



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 168293

TO: Jeffrey Parkin
Location: REM-3D39/3C18
Art Unit: 1648
Wednesday, August 03, 2005
Case Serial Number: 10/066506

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Parkin,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

This Page Blank (uspto)

STIC-Biotech/ChemLib

160293

From: Parkin, Jeffrey
Sent: Sunday, July 24, 2005 5:36 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 10/066,506

Please search **SEQ ID NOS.: 1, 3, 5, and 7** from the aforementioned application (U.S. Serial No. 10/066,506) v. all relevant databases. Place results on both paper and electronic format (i.e., disk, e-mail, etc.).

Thanks!

JSP
AU 1648
REM 3D39
2-0908

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 9678.51 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-1

Perfect score: 2252

Sequence: 1 atggcgcttacaggaaatt.....tatatggataacggaattc 2252

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2252	100.0	2252	6	AX717713	Sequence
2	1506	66.9	2039	6	AX717723	Sequence
3	1501.4	66.7	11742	14	EBORNA	L11365 Zaire Ebola
4	1499.8	66.6	2298	6	AR404895	Sequence
5	1499.8	66.6	2298	6	AX092102	Sequence
6	1499.4	66.1	2406	14	EVU131033	U31033 Zaire Ebola
7	1489.4	66.1	2408	14	EVU232187	U23187 Zaire Ebola
8	1489.4	66.1	18959	14	AF086833	AF086833 Zaire Ebo
9	1489.4	66.1	18959	14	AF272001	AF272001 Zaire Ebo
10	1489.4	66.1	18959	14	AY142960	AY142960 Zaire Ebo
11	1487.8	66.1	2408	14	EVU81161	U81161 Zaire Ebola
12	1486.2	66.0	18960	14	AF499101	AF499101 Zaire Ebo
13	1464.6	65.0	18961	14	AY354458	AY354458 Zaire ebo
14	1452.6	64.5	2408	14	EVU28077	U28077 Zaire Ebola
15	1446.2	64.2	2446	14	EVU77384	U77384 Zaire Ebola
16	1444.6	64.1	2173	14	AY058898	AY058898 Zaire Ebo
17	1435	63.7	2173	14	AY526098	AY526098 Zaire ebo
18	1435	63.7	2173	14	AY526099	AY526099 Zaire ebo
19	1435	63.7	2173	14	AY526100	AY526100 Zaire ebo

20	1435	63.7	2173	14	AY526102	AY526102 Zaire ebo
21	1435	63.7	2173	14	AY526104	AY526104 Zaire ebo
22	1435	63.7	2173	14	AY526105	AY526105 Zaire ebo
23	1433.4	63.7	2173	14	AY526101	AY526101 Zaire ebo
24	753	33.4	2051	6	AX717721	Sequence
25	746.6	33.2	2046	6	AX717719	Sequence
26	742.2	33.0	11460	6	AR279677	Sequence
27	742.2	33.0	19104	14	MVRBPCYC	Z12332 Marburg vir
28	740.6	32.9	19112	14	AY430366	AY430366 Lake Vict
29	740.6	32.9	19113	14	AY430365	AY430365 Lake Vict
30	685.8	30.5	2046	14	AF005735	AF005735 Marburg v
31	685.8	30.5	2948	14	MAVSPAA	X68493 Marburg Vir
32	685.8	30.5	19112	14	MVIRPR	Z29337 Marburg vir
33	675	30.0	19151	14	AY558025	AY358025 Marburg v
34	673	29.9	2046	14	AF005733	AF005733 Marburg v
35	548.2	24.3	2046	6	AX717725	AX717725 Sequence
36	546.6	24.3	2046	6	AX717717	AX717717 Sequence
37	543.4	24.1	2046	14	AF005734	AF005734 Marburg v
38	534.2	23.7	2408	14	EVU28006	U28006 Cote d'Ivoi
39	466.2	20.7	18890	14	AB050936	AB050936 Reston Eb
40	464.6	20.6	2360	14	EVU23152	U23152 Reston Ebol
41	463	20.6	2233	14	AF034645	AF034645 Ebola vir
42	461.4	20.5	18891	14	AF522874	AF522874 Reston Eb
43	459.8	20.4	2033	14	EVU23416	U23416 Reston Ebol
44	456.6	20.3	2033	14	EVU23417	U23417 Reston Ebol
45	428.2	19.0	2261	14	AY344234	AY344234 Sudan ebo

ALIGNMENTS

RESULT 1
AX717713
LOCUS AX717713 2252 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 1 from Patent WO02079239.
ACCESSION AX717713
VERSION AX717713.1 GI:29890725
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 1 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2252
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 1 and Marburg virus strain Musoke Glycoprotein 2"

ORIGIN	Query Match	100.0%;	Score 2252;	DB 6;	Length 2252;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2252;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATCTTTT	60		
Db	1	ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATCTTTT	60		
Qy	61	CTTTGGTAAATATTCCTTTTCCAAAGAACATTTCCATCCCATCTTGAGTCATCCCAAT	120		
Db	61	CTTTGGTAAATATTCCTTTTCCAAAGAACATTTCCATCCCATCTTGAGTCATCCCAAT	120		
Qy	121	AGCATATTACAGGTAGTAGTGTGCGACAACTAGTTTGTGTCGACAACTGTCATCCACA	180		
Db	121	AGCATATTACAGGTAGTAGTGTGCGACAACTAGTTTGTGTCGACAACTGTCATCCACA	180		
Qy	181	AATCAATTGATGATCAGTTGGACTGATCTCGAAGGGAAATGGAGTGGCAACTGACGTCCA	240		

181 AATCAATTGAGATCAGTTGGATGGAATCTCGAAGGAATGGAGTGGCAACTGACGTGCCA 240
241 TCTGCAACTAAAGAGATGGGCTTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA 300
241 TCTGCAACTAAAGAGATGGGCTTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA 300
301 GCTGTGAAATGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTTGACCGGAGTGAG 360
301 GCTGTGAAATGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTTGACCGGAGTGAG 360
361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGTCGGGTATGTGACAAA 420
361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGTCGGGTATGTGACAAA 420
421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCCCTCCATAAAGAGGGTCTTCTTC 480
421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCCCTCCATAAAGAGGGTCTTCTTC 480
481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTGCTGAAGTGTG 540
481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTGCTGAAGTGTG 540
541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAAGGACTTCTCAGCTCACACCCCTTGAGA 600
541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAAGGACTTCTCAGCTCACACCCCTTGAGA 600
601 GAGCGGTCAATGCAACGGGAGACCGTCTAGTGGCTACTATTCTACCAAAATAGATAT 660
601 GAGCGGTCAATGCAACGGGAGACCGTCTAGTGGCTACTATTCTACCAAAATAGATAT 660
661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTGACC 720
661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTGACC 720
721 TAGCTCCAACTTGAATCAAGATTCAACCAAGTTCGCTCAGCTGAATGAGACAATA 780
721 TAGCTCCAACTTGAATCAAGATTCAACCAAGTTCGCTCAGCTGAATGAGACAATA 780
781 TATACAAGTGGGAAAGAGGCAATACCGGGAACCAATTTGGAAGGTCAACCCCGAA 840
781 TATACAAGTGGGAAAGAGGCAATACCGGGAACCAATTTGGAAGGTCAACCCCGAA 840
841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAACCTCACTAGAAAA 900
841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAACCTCACTAGAAAA 900
901 ATTGCAAGTGAAGTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAACATCAAGTGT 960
901 ATTGCAAGTGAAGTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAACATCAAGTGT 960
961 CAGAGTCCGGCGGAATCTTCCGACCCAGGACCAACAACTGAGACCAAAA 1020
961 CAGAGTCCGGCGGAATCTTCCGACCCAGGACCAACAACTGAGACCAAAA 1020
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAGGCT 1080
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAGGCT 1080
1081 GCAGTGTGCAATCAACACCTTGGCAAACTCTCCACGAGTCCCAATCCCTCAACCC 1140
1081 GCAGTGTGCAATCAACACCTTGGCAAACTCTCCACGAGTCCCAATCCCTCAACCC 1140
1141 AAACCAAGTCCGGACAAAGCCCAATATACCCGTGTATAAATCTTGACATCTCTGAG 1200
1141 AAACCAAGTCCGGACAAAGCCCAATATACCCGTGTATAAATCTTGACATCTCTGAG 1200
1201 GCAACTCAAGTTGAACAACTCACCCGAGAACAGACAGACAGACAGCTCCGACACT 1260
1201 GCAACTCAAGTTGAACAACTCACCCGAGAACAGACAGACAGACAGCTCCGACACT 1260
1261 CCTCTTGCCACGACCGGACCGGACCCCAAGAGAGAACCAACCAACGAGGAGAGGC 1320

1261 CCTCTGCCACGACCGCAGCCGACCCCAAAAGCAGAGAAACACCAACGACGAGAGGC 1320
1321 ACTGACTTCTGGAGCCCGCCACACAAAGTCCCAAAACACAGAGAGACCGCTGGC 1380
1321 ACTGACTTCTGGAGCCCGCCACACAAAGTCCCAAAACACAGAGAGACCGCTGGC 1380
1381 AACAAACAACATCATACCAAGATACCGGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1440
1381 AACAAACAACATCATACCAAGATACCGGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1440
1441 TTAATTACCAATACATTTGCTGGAGTCCAGAGCTGATCAGCGGGAAGAGAACTCGA 1500
1441 TTAATTACCAATACATTTGCTGGAGTCCAGAGCTGATCAGCGGGAAGAGAACTCGA 1500
1501 CGATCGATCTCTGGAGGAAAGGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCT 1560
1501 CGATCGATCTCTGGAGGAAAGGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCT 1560
1561 CCAATTGATTTGAGCCAGTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCT 1620
1561 CCAATTGATTTGAGCCAGTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCT 1620
1621 GGTGCTCGGCTGAGGAAGATCAACATGCTCTCCCAATTTAGTTAACTTTATCTTAT 1680
1621 GGTGCTCGGCTGAGGAAGATCAACATGCTCTCCCAATTTAGTTAACTTTATCTTAT 1680
1681 TTTCTTAATATAAATGAGAAACAATGCTCTCTGAGAAAATGAGAATGATGTGATGCA 1740
1681 TTTCTTAATATAAATGAGAAACAATGCTCTCTGAGAAAATGAGAATGATGTGATGCA 1740
1741 GAGTTAAGAAATTTGAGCGTTTCAAGAGATGACCTGGCGGAGGCTCAGTTGGATACCG 1800
1741 GAGTTAAGAAATTTGAGCGTTTCAAGAGATGACCTGGCGGAGGCTCAGTTGGATACCG 1800
1801 TTTTGGCCCTGGAAATTTGAAGGACTTTACACTGCTGTTTAAATTTAAAAATCAAAACAT 1860
1801 TTTTGGCCCTGGAAATTTGAAGGACTTTACACTGCTGTTTAAATTTAAAAATCAAAACAT 1860
1861 TTGGTCTCAGGTTGAGCGCTTAGCCAATCAAACTGCCAAATCTTGGAACTCTTATG 1920
1861 TTGGTCTCAGGTTGAGCGCTTAGCCAATCAAACTGCCAAATCTTGGAACTCTTATG 1920
1921 AGAGTCAACAACCTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTCTA 1980
1921 AGAGTCAACAACCTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTCTA 1980
1981 CTCACAAGATGGGAGGAAACATGCAAGTCTTGGACCTGATTTGCAATCGGATAGAA 2040
1981 CTCACAAGATGGGAGGAAACATGCAAGTCTTGGACCTGATTTGCAATCGGATAGAA 2040
2041 GACTTGTCCAAAATATTTTCAGAGCAAAATTGACCAAAATTTAAAAAGGAGAAACAAAAGAG 2100
2041 GACTTGTCCAAAATATTTTCAGAGCAAAATTGACCAAAATTTAAAAAGGAGAAACAAAAGAG 2100
2101 GGGACTGTGTGGGTCTGGGTGTTAAATGGTGGACATCCGACTGGGGTCTTCTTACTAAC 2160
2101 GGGACTGTGTGGGTCTGGGTGTTAAATGGTGGACATCCGACTGGGGTCTTCTTACTAAC 2160
2161 TTGGGCATTTTGTCTACTATTATCCATAGTCTGTTGATGCTCTATCCTGATTTTGTCTG 2220
2161 TTGGGCATTTTGTCTACTATTATCCATAGTCTGTTGATGCTCTATCCTGATTTTGTCTG 2220
2221 ATCTTTTACTAAATATATCGGATAACGGAATTC 2252
2221 ATCTTTTACTAAATATATCGGATAACGGAATTC 2252

RESULT 2
AX717723
LOCUS
DEFINITION
ACCESSION
VERSION

AX717723 2039 bp DNA linear PAT 15-APR-2003
Sequence 11 from Patent WO02079239.
AX717723 GI:29890730

KEYWORDS Zaire ebolavirus (ZEBOV)
SOURCE Zaire ebolavirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 11 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
1..2039
/organism="Zaire ebolavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:186538"
/note="chimeric molecule between Ebola virus Glycoprotein
1 and Ebola virus Glycoprotein 2"
ORIGIN
Query Match 66.9%; Score 1506; DB 6; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 60
Db 1 ATGGGCGTTACAGGAATATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 60
Qy 61 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
Db 61 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
Qy 121 AGCATTACAGGTTAGTGATGTGACAACTAGTTTGTGCTGACAAACTGTGATCCACA 180
Db 121 AGCATTACAGGTTAGTGATGTGACAACTAGTTTGTGCTGACAAACTGTGATCCACA 180
Qy 181 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGAAATGAGTGGCAACTGAGTGCCA 240
Db 181 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGAAATGAGTGGCAACTGAGTGCCA 240
Qy 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCCACCAAAAGGTGGTCAATATGAA 300
Db 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCCACCAAAAGGTGGTCAATATGAA 300
Qy 301 GCTGTGTAATGGGTGAAAACTGCTACAACTTTGAAATCAAAAAAAGGTGAGTGGAG 360
Db 301 GCTGTGTAATGGGTGAAAACTGCTACAACTTTGAAATCAAAAAAAGGTGAGTGGAG 360
Qy 361 TGTCTACAGAGCGCCAGAGCGGATTCGGGGCTTCCCGGTGCGCGTATGTGCACAAA 420
Db 361 TGTCTACAGAGCGCCAGAGCGGATTCGGGGCTTCCCGGTGCGCGTATGTGCACAAA 420
Qy 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAGAGGGTCTTTCTTC 480
Db 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAGAGGGTCTTTCTTC 480
Qy 481 CTGTATGATCGACTTGTTCACAGTTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 540
Db 481 CTGTATGATCGACTTGTTCACAGTTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 540
Qy 541 GTTGCAATTTCTGATCTGCTCCCAAGCTAAGAGGACTTTTCAGCTACACCCCTTGAGA 600
Db 541 GTTGCAATTTCTGATCTGCTCCCAAGCTAAGAGGACTTTTCAGCTACACCCCTTGAGA 600
Qy 601 GAGCCGGTCAATGCAACCGAGGACCGGTCTAGTGGCTACTATTCTACCAATATAGATAT 660
Db 601 GAGCCGGTCAATGCAACCGAGGACCGGTCTAGTGGCTACTATTCTACCAATATAGATAT 660
Qy 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTTGTCGAGGTTGACAAATTGACC 720
Db 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTTGTCGAGGTTGACAAATTGACC 720
Qy 721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780

Db 721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Qy 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAACATAATTTGGAAGGTCAACCCGAA 840
Db 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAACATAATTTGGAAGGTCAACCCGAA 840
Qy 841 ATTGATACAAATCGGGAGTGGGCTTCTGGGAACTAATAAAAAACCTCAGTAAAA 900
Db 841 ATTGATACAAATCGGGAGTGGGCTTCTGGGAACTAATAAAAAACCTCAGTAAAA 900
Qy 901 ATTGCGAGTCAAGAGTGTCTTTTCCAGAGTGTATCAAAACGAGGCAAAAAACATCAGTGT 960
Db 901 ATTGCGAGTCAAGAGTGTCTTTTCCAGAGTGTATCAAAACGAGGCAAAAAACATCAGTGT 960
Qy 961 CAGAGTCCGGCGGAACTTTTCCGAGCCAGGAGCAACCAACAACTGAAGACCAAAA 1020
Db 961 CAGAGTCCGGCGGAACTTTTCCGAGCCAGGAGCAACCAACAACTGAAGACCAAAA 1020
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTTCAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTTCAGTGCACAGTCAAGGAAGGAAAGCT 1080
Qy 1081 GCAGTGTGCAATCTTAACAACCTTGGCCCAATCTCCACGAGTCCCAATCCCTCACAAC 1140
Db 1081 GCAGTGTGCAATCTTAACAACCTTGGCCCAATCTCCACGAGTCCCAATCCCTCACAAC 1140
Qy 1141 AAACAGGTCGGGACAAACAGCAGCCCAATATACACCCGTGTATAAATTTGACATCTCTGAG 1200
Db 1141 AAACAGGTCGGGACAAACAGCAGCCCAATATATACACCCGTGTATAAATTTGACATCTCTGAG 1200
Qy 1201 GCAACTCAAGTTGAAACAACTACCGCAGACAGACAGACAGCAGCAGCTCCGACACT 1260
Db 1201 GCAACTCAAGTTGAAACAACTACCGCAGACAGACAGACAGCAGCAGCTCCGACACT 1260
Qy 1261 CCCTCTGCCAGCAGCGGAGCCGAGCCCAAAAGCAGAGAAACCAACACGAGCAAGAGC 1320
Db 1261 CCCTCTGCCAGCAGCGGAGCCGAGCCCAAAAGCAGAGAAACCAACACGAGCAAGAGC 1320
Qy 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACGAGAGACCGCTGGC 1380
Db 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACGAGAGACCGCTGGC 1380
Qy 1381 AACCAACAACTCATACCAAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
Db 1381 AACCAACAACTCATACCAAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
Qy 1441 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
Db 1441 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
Qy 1501 CGATCG 1506
Db 1501 CGATCG 1506
RESULT 3
EBORNA 11742 bp ss-RNA linear VRL 28-AUG-2002
LOCUS Zaire Ebola virus nucleoprotein, polymerase complex protein (vp35),
DEFINITION matrix protein (vp40), glycoprotein (gp), minor nucleoprotein
(vp30), and membrane-associated structural protein (vp24)s,
complete cds; and polymerase (POLYMERASE), partial cds.
ACCESSION L11365
VERSION L11365
KEYWORDS L11365.1 GI:2522270
ORGANISM Zaire ebolavirus (ZEBOV)
SOURCE Viruses; ssRNA negative-strand viruses; Mononegavirales;
REFERENCE 1 (bases 1 to 11742)
AUTHORS Sanchez, A., Kiley, M.P., Holloway, B.P., McCormick, J.B. and
Auperin, D.D.
TITLE The nucleoprotein gene of Ebola virus: cloning, sequencing, and in

misc_feature	7792..7869	/gene="gp"	
misc_feature	/note="encodes immunosuppressive motif"		
	7987..8055	/gene="gp"	
	/note="encodes transmembrane domain"		
gene	8291..9741	/gene="vp30"	
mRNA	8291..9741	/gene="vp30"	
	/product="minor nucleoprotein"		
	/note="putative"		
	/citation=[2]		
CDS	8511..9293	/gene="vp30"	
	/note="putative"		
	/citation=[2]		
	/codon_start=1		
	/product="minor nucleoprotein"		
	/protein_id="AAB81005.1"		
	/db_xref="GI:290393"		
	/translation="MEASYSERGRPRARQHSRDHHRARSSSRENYRGEYRQSR		
	ASQVRPVPFHKRVPEIPTVPAPKDIPTLTKGFLCDSPFKKHOLESLDRELL		
	LIARTCGSVQOLNITAPKDSRLANPTADDFQOBEGPKITLLTLIKTAHWRQDIR		
	TIEDSKRALITLCNVRKESKLSLLCETHLRREGLGQDQAPVLEVYQRLHSDK		
	GGSEFALUQWQDQSLIMFIFALNIALQLPCESSAALFQG"		
misc_RNA	complement(9742..9883)		
	/product="intergenic region"		
	/citation=[2]		
	/function="separates vp30 gene and vp24 gene"		
	/evidence=experimental		
gene	9884..11518	/gene="vp24"	
mRNA	9884..11518	/gene="vp24"	
	/product="membrane-associated structural protein"		
	/citation=[2]		
	/evidence=experimental		
Query Match 66.7%; Score 1501.4; DB 14; Length 11742;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGGCGGTACAGGAATATGCAATTACCTCGTGAATCGAATTCAGAGACATCAATCTTTT	60
Db	6041	ATGGCGGTACAGGAATATGCAATTACCTCGTGAATTCAGAGACATCAATCTTTT	6100
Qy	61	CTTTGGGTAATATCTCTTTCCAGAGACATTTTCCATCCGATGGAGTCAACCAAT	120
Db	6101	CTTTGGGTAATATCTCTTTCCAGAGACATTTTCCATCCGATGGAGTCAACCAAT	6160
Qy	121	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAAACTGTCTATCCACA	180
Db	6161	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAAACTGTCTATCCACA	6220
Qy	181	AATCAATTGAGATCAGTTGGATGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCA	240
Db	6221	AATCAATTGAGATCAGTTGGATGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCA	6280
Qy	241	TCTGCAATTAAGATGGGCTTCAGGTCGGGTGCCCAACCAAGGTGGTCAATATGAA	300
Db	6281	TCTGCAATTAAGATGGGCTTCAGGTCGGGTGCCCAACCAAGGTGGTCAATATGAA	6340
Qy	301	GCTGTGTAATGGGCTGAAACTGTCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG	360
Db	6341	GCTGTGTAATGGGCTGAAACTGTCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG	6400
Qy	361	TGTCTACAGACGCCACAGACGGGATTCGGGCTTTCCCGGTGCGGTATGTGCACAAA	420
Db	6401	TGTCTACAGACGCCACAGACGGGATTCGGGCTTTCCCGGTGCGGTATGTGCACAAA	6460
Qy	421	GTATCAGGAACGGGACCGGTGTCGGGACATTTGCTTCCATTAAGAGGGTGCTTTCTTC	480

Db	6461	GTATCAGGAACGGGACCGGTGTCGGGACATTTGCTTCCATTAAGAGGGTGCTTTCTTC	6520
Qy	481	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC	540
Db	6521	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC	6580
Qy	541	GTTGCATTTCTGATCTACTGCCCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	600
Db	6581	GTTGCATTTCTGATCTACTGCCCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	6640
Qy	601	GAGCCGGTCAATGCAACGAGGAGCCCGTCTAGTGGCTACTATTCTTACCACAAATAGATAT	660
Db	6641	GAGCCGGTCAATGCAACGAGGAGCCCGTCTAGTGGCTACTATTCTTACCACAAATAGATAT	6700
Qy	661	CAGCTACCGGTTTTGGAAACAATGAGACAGATGACTTTGTTTCGAGGTTGACAAATTTGACC	720
Db	6701	CAGCTACCGGTTTTGGAAACAATGAGACAGATGACTTTGTTTCGAGGTTGACAAATTTGACC	6760
Qy	721	TACGTCCAACCTTGAATCAAGATTTCACACCACAGTTTCTGCTCCAGCTGAATGAGACATA	780
Db	6761	TACGTCCAACCTTGAATCAAGATTTCACACCACAGTTTCTGCTCCAGCTGAATGAGACATA	6820
Qy	781	TATCAAGTGGGAAAGGAGCAATACACAGGGAACCTTAATTTGGAAGGTCAACCCGAA	840
Db	6821	TATCAAGTGGGAAAGGAGCAATACACAGGGAACCTTAATTTGGAAGGTCAACCCGAA	6880
Qy	841	ATTGATACAACATCGGGAGTGGGCTTCTGGGAACTAAAAAACCTCCTCATTAGAAAA	900
Db	6881	ATTGATACAACATCGGGAGTGGGCTTCTGGGAACTAAAAAACCTCCTCATTAGAAAA	6940
Qy	901	ATTGCGAGTGAAGAGTTGTCTTTTTCAGAGTTGTATCAACGAGGACCAAAACATCATGCTGT	960
Db	6941	ATTGCGAGTGAAGAGTTGTCTTTTTCAGAGTTGTATCAACGAGGACCAAAACATCATGCTGT	7000
Qy	961	CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACCAACAACTGAAGACCAACAA	1020
Db	7001	CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACCAACAACTGAAGACCAACAA	7060
Qy	1021	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT	1080
Db	7061	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT	7120
Qy	1081	GCAGTGTGCATCTAACAAACCTTTGCCACATCTCCAGAGTCCCAATCCCTCAACACC	1140
Db	7121	GCAGTGTGCATCTAACAAACCTTTGCCACATCTCCAGAGTCCCAATCCCTCAACACC	7180
Qy	1141	AAACAGGTCGGGCAACAGCACCCATTAACACCCGTTGATAAACTTGACATCTCTGAG	1200
Db	7181	AAACAGGTCGGGCAACAGCACCCATTAACACCCGTTGATAAACTTGACATCTCTGAG	7240
Qy	1201	GCAACTCAAGTTGAACAACTATCACCAGAACAGAACAGACAGCAAGCTCCGACACT	1260
Db	7241	GCAACTCAAGTTGAACAACTATCACCAGAACAGAACAGACAGCAAGCTCCGACACT	7300
Qy	1261	CCCTCTGCGACGACCGGACCGGACCCGCCAAAAGAGAGAAACCAACAGGAGAGC	1320
Db	7301	CCCTCTGCGACGACCGGACCGGACCCGCCAAAAGAGAGAAACCAACAGGAGAGC	7360
Qy	1321	ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACCAACAGGAGAGCAGCGTGGC	1380
Db	7361	ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACCAACAGGAGAGCAGCGTGGC	7420
Qy	1381	AACAAACACTCTATCACCAGATACCGGAGAGAGAGTGCCAGCAGCGGAGAGCTAGGC	1440
Db	7421	AACAAACACTCTATCACCAGATACCGGAGAGAGAGTGCCAGCAGCGGAGAGCTAGGC	7480
Qy	1441	TTAATTACCAATATCTATTGCTGGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA	1500
Db	7481	TTAATTACCAATATCTATTGCTGGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA	7540
Qy	1501	CGA 1503	
Db	7541	AGA 7543	

```
RESULT 4
AR404895
LOCUS AR404895 2298 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6630144.
ACCESSION AR404895
VERSION AR404895.1 GI:40153661
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2298)
AUTHORS Hart, M.K., Wilson, J.A. and Schmaljohn, A.L.
TITLE Monoclonal antibodies to Ebola glycoprotein
JOURNAL Patent: US 6630144-A 1 07-OCT-2003;
FEATURES
source
1..2298
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 66.6%; Score 1499.8; DB 6; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGGGGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 60
DB |||||||
QY 59 ATGGGGGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATCTTT 118
DB |||||||
QY 61 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT 120
DB |||||||
QY 119 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT 178
DB |||||||
QY 121 AGCACAATTACAGGTTAGTATGATCGCAAACTAGTTTGTGTCGACAAATGTCATCCCA 180
DB |||||||
QY 179 AGCACATTACAGGTTAGTATGATCGCAAACTAGTTTGTGTCGACAAATGTCATCCCA 238
DB |||||||
QY 181 AATCAATTACAGGTCAGTTGGATCTCGAAGGAATGGAATGGAATGGAATGGAATGGAAT 240
DB |||||||
QY 239 AATCAATTGAGATCAGTTGGATCTCGAAGGAATGGAATGGAATGGAATGGAATGGAAT 298
DB |||||||
QY 241 TTGCAACTAAAGATGGGGCTTCAGGTCGGGTGCCACCAAGGTGGTCAATTATGAA 300
DB |||||||
QY 299 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGCCACCAAGGTGGTCAATTATGAA 358
DB |||||||
QY 301 GCTGGTGAATGGGCTGAAACTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGG 360
DB |||||||
QY 359 GCTGGTGAATGGGCTGAAACTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGG 418
DB |||||||
QY 361 TGTCTACCAAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTCGCCGATATGTCACAAA 420
DB |||||||
QY 419 TGTCTACCAAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTCGCCGATATGTCACAAA 478
DB |||||||
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCCCTTCCATAAAGAGGGTCTTTCTTC 480
DB |||||||
QY 479 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCCCTTCCATAAAGAGGGTCTTTCTTC 538
DB |||||||
QY 481 CTGTATGATCGACTTGTCTCCAGTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 540
DB |||||||
QY 539 CTGTATGATCGACTTGTCTCCAGTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 598
DB |||||||
QY 541 GTTGCATTTCTGATATGCCCCAAGCTAAGAGGACTTCTTCAGCTCAGCCCTTGAGA 600
DB |||||||
QY 599 GTTGCATTTCTGATATGCCCCAAGCTAAGAGGACTTCTTCAGCTCAGCCCTTGAGA 658
DB |||||||
QY 601 GAGCCGGTCAATGCAACGAGGACCGGTCTAGTGGCTACTATTCTACCAAAATAGATAT 660
DB |||||||
QY 659 GAGCCGGTCAATGCAACGAGGACCGGTCTAGTGGCTACTATTCTACCAAAATAGATAT 718
DB |||||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTCGACAAATTTGACC 720
DB |||||||
QY 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTTGACAAATTTGACC 778
DB |||||||

RESULT 5
AX092102 2298 bp DNA linear PAT 29-AUG-2002
LOCUS AX092102
DEFINITION Sequence 1 from Patent WO0116183.
ACCESSION AX092102
VERSION AX092102.1 GI:13444354
KEYWORDS
SOURCE Zaire ebolavirus (ZBOV)
ORGANISM Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE
1
AUTHORS Hart, M.K., Wilson, J.A. and Schmaljohn, A.L.
TITLE Monoclonal antibodies and vaccines against epitopes on the ebola
virus glycoprotein
```

JOURNAL Patent: WO 0116193-A 1 08-MAR-2001;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
source 1..2298
/organism="Zaire ebolavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:186538"

ORIGIN

Query Match 66.6%; Score 1499.8; DB 6; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAAATATTCAGTTACCTCGTGAATCGATTCAAGAGGACATCATCTTT 60
DB 59 ATGGCGGTTACAGGAAATATTCAGTTACCTCGTGAATCGATTCAAGAGGACATCATCTTT 118
QY 61 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 120
DB 119 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 178
QY 121 AGCATTATACAGGTTAGTGTGACAAATAGTTTGTGTCGACAACTGTGTCATCCACA 180
DB 179 AGCATTATACAGGTTAGTGTGACAAATAGTTTGTGTCGACAACTGTGTCATCCACA 238
QY 181 AATCAATTGAGATCAGTTGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
DB 239 AATCAATTGAGATCAGTTGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
QY 241 TCTGCAATTAAGATGGGCTTCAGGTCGGGTGCCACCAAAAGGTGCTCAATTTATGAA 300
DB 299 TCTGCAATTAAGATGGGCTTCAGGTCGGGTGCCACCAAAAGGTGCTCAATTTATGAA 358
QY 301 GCTGTGAATGGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGAGTGAG 360
DB 359 GCTGTGAATGGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGAGTGAG 418
QY 361 TGTCTACAGACGCCACGACAGCGGATTCGGGGCTTCCCGGTCGGGTATGTGACAAAA 420
DB 419 TGTCTACAGACGCCACGACAGCGGATTCGGGGCTTCCCGGTCGGGTATGTGACAAAA 478
QY 421 GTATCAGGAACGGGACCGGTGTCGGGACACTTTCCTTCCATAAGAGGGTCTTCTTC 480
DB 479 GTATCAGGAACGGGACCGGTGTCGGGACACTTTCCTTCCATAAGAGGGTCTTCTTC 538
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGC 540
DB 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGC 598
QY 541 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGTCTACACCCCTTGAGA 600
DB 599 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGTCTACACCCCTTGAGA 658
QY 601 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGGCTACTATTCTACCAATAGATAT 660
DB 659 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGGCTACTATTCTACCAATAGATAT 718
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
DB 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 778
QY 721 TAGCTCCAACTTGAATCAAGATTCTACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 779 TAGCTCCAACTTGAATCAAGATTCTACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 838
QY 781 TATACAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTGGAAGTCAACCCGAA 840
DB 839 TATACAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTGGAAGTCAACCCGAA 898
QY 841 ATTGTATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAAACTCACTAGAAAA 900
DB 899 ATTGTATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAAACTCACTAGAAAA 958

QY 901 ATTCCGAGTGAAGAGTTGTCTTTTTCAGAGTTGTATCAAAACGAGCCAAAACATCAGTGT 960
DB 959 ATTCCGAGTGAAGAGTTGTCTTTTTCAGAGTTGTATCAAAACGAGCCAAAACATCAGTGT 1018
QY 961 CAGAGTCCGGCGGAACTTCTTCGAGCCAGGAGCAACAACTGTAAGACCAAAA 1020
DB 1019 CAGAGTCCGGCGGAACTTCTTCGAGCCAGGAGCAACAACTGTAAGACCAAAA 1078
QY 1021 ATCATGGCTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT 1080
DB 1079 ATCATGGCTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT 1138
QY 1081 GCAGTGTGCGATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
DB 1139 GCAGTGTGCGATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1198
QY 1141 AAACAGGTCGGGAAACAGCACCATAATACACCGTGTATAAACTTGACATCTCTGAG 1200
DB 1199 AAACAGGTCGGGAAACAGCACCATAATACACCGTGTATAAACTTGACATCTCTGAG 1258
QY 1201 GCAACTCAAGTTGAACAACTCACCAGAACAGAACAAAGCAGACAGCTCCGACACT 1260
DB 1259 GCAACTCAAGTTGAACAACTCACCAGAACAGAACAAAGCAGACAGCTCCGACACT 1318
QY 1261 CCCTCTGCCAGCAGCGGAGCCGACCCCAAAAGCAGAGAACACACGAGCAAGGC 1320
DB 1319 CCCTCTGCCAGCAGCGGAGCCGACCCCAAAAGCAGAGAACACACGAGCAAGGC 1378
QY 1321 ACTGACTTCTGGAGCCCGCCACCAACAACTGCCCAAAACACAGAGACCGCTGGC 1380
DB 1379 ACTGACTTCTGGAGCCCGCCACCAACAACTGCCCAAAACACAGAGACCGCTGGC 1438
QY 1381 AACAACTCACTCATCAACCAAGATACCGGAGAAAGAGTGCCAGAGCGGGAAGCTAGGC 1440
DB 1439 AACAACTCACTCATCAACCAAGATACCGGAGAAAGAGTGCCAGAGCGGGAAGCTAGGC 1498
QY 1441 TTAATTAACCAATATCTTGTGGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA 1500
DB 1499 TTAATTAACCAATATCTTGTGGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA 1558
QY 1501 CGA 1503
DB 1559 AGA 1561

RESULT 6
EVU31033
LOCUS
DEFINITION Zaire Ebola virus envelope glycoprotein (GP) gene, complete cds,
secreted small glycoprotein (sgp) gene, complete cds.
ACCESSION U31033 S59050
VERSION U31033.1 GI:1141778
KEYWORDS
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (sites)
AUTHORS Volchkov V.E., Becker S., Volchkova V.A., Ternovoj, V.A.,
Kotov A.N., Netesov, S.V. and Klenk, H.D.
TITLE GP mRNA of Ebola virus is edited by the Ebola virus polymerase and
by T7 and vaccinia virus polymerases
JOURNAL Virology 214 (2), 421-430 (1995)
MEDLINE 96130181
PUBMED 8553543
REFERENCE 2 (sites)
AUTHORS Volchkov V.E., Blinov V.M., Kotov, A., Chepurnov, A. and Netesov, S.
TITLE The full-length nucleotide sequence of the Ebola virus
JOURNAL Thesis (1993)
REFERENCE 3 (sites)
AUTHORS Volchkov V.E., Blinov V.M. and Netesov, S.V.
TITLE The envelope glycoprotein of Ebola virus contains an
immunosuppressive-like domain similar to oncogenic retroviruses


```
JOURNAL FEBS Lett. 305 (3), 181-184 (1992)
MEDLINE 93231342
PUBMED 1299611
REFERENCE 4 (bases 1 to 2406)
AUTHORS Volchkov, V.E.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1995) Viktor E. Volchkov, Institut fuer
Virologie, Philipps-Universitaet, Robert-Koch-Str.17, Marburg
35037, Germany
COMMENT On Dec 31, 1995 this sequence version replaced gi:299674.
FEATURES Location/Qualifiers
     source
       1..2406
         /organism="Zaire ebolavirus"
         /mol_type="genomic RNA"
         /strain="Zaire"
         /db_xref="taxon:186538"
     CDS
       join(140..1162,1162..2169)
       /note="Synthesis of the full length EBO GP occurs from
       edited form of GP mRNA; one additional A residue inserts
       in EBO GP editing site."
       /citation=[3]
       /citation=[1]
       /citation=[2]
       /codon_start=1
       /evidence=experimental
       /product="envelope glycoprotein"
       /protein_id="AAA96744.1"
       /db_xref="GI:1141779"
       /translation="MGVIGILQLPRDRFKRTSFFLWVILFQRTSIPLGVIHNSLQ
       VSDVKLVCRDKLSNQURSVGLNLENGVAIDVPSATKRWGFRSGVPPKVVNTEAG
       EWAENCYNIEIKKDPGSECLPAAPDGIRGPRCYRVHKVSGTGPCAGDPAFHKEGAF
       LYDRLASTVIYRGTTFAEGVAFLILPOAKKDFSSHPLREPNVATEDPSSGYSTTI
       RYQATGFGNETYELFVDNLTVOLESRTPOFLQLNETIYTSKRSNTTGLKIWK
       VNPEIDTITGEWAFETKTSLEKFAVKSCLSQLYQTEPKTSVVRRELLPTQGPQ
       QUKTKSMASENSAMVQHSQGREAAVSHLTILATISPSQSLTKGPDNSHNTP
       VYKLDISEATQVEQHRRTDNDSTADTSPATTAAGPPKAEINTSKSTDFLDPATTT
       SPQNSHTANNTHQDDTGESASGKGLGINTIAGVAGLITGRRTREAIVNAQ
       PCKNPLNHWYTTQDEGAATGLAWIPYFGPAEGIVIEGLMHNQDLGICGLQLANETT
       QALQLPLRATTFEISLIRKAIIDFLQRWGTCHILGPDCIIEPHDWTNITDKID
       QIIEHVDKTLDPQGDNDNWTGWRQWIPAGIVGTGVIIAVIALFCICKFVP"
       140..1234
       /gene="SGP"
       140..1234
       /gene="sgp"
       /note="synthesis on non-edited form of EBO GP mRNA"
       /species="
       /citation=[2]
       /codon_start=1
       /evidence=experimental
       /product="secreted small glycoprotein"
       /protein_id="AAA96745.1"
       /db_xref="GI:1141780"
       /translation="MGVIGILQLPRDRFKRTSFFLWVILFQRTSIPLGVIHNSLQ
       VSDVKLVCRDKLSNQURSVGLNLENGVAIDVPSATKRWGFRSGVPPKVVNTEAG
       EWAENCYNIEIKKDPGSECLPAAPDGIRGPRCYRVHKVSGTGPCAGDPAFHKEGAF
       LYDRLASTVIYRGTTFAEGVAFLILPOAKKDFSSHPLREPNVATEDPSSGYSTTI
       RYQATGFGNETYELFVDNLTVOLESRTPOFLQLNETIYTSKRSNTTGLKIWK
       VNPEIDTITGEWAFETKTSLEKFAVKSCLSQLYQTEPKTSVVRRELLPTQGPQ
       QUKTKSWLQKTLPLQWFKCTVKEGLQCR1"
       1019..1025
       /gene="sgp"
       /note="RNA editing site; synthesis of full length GP
       proceeds via RNA editing; one additional A residue inserts
       during transcription of the Ebola virus GP mRNA in region;
       contains 7 consecutive A residues."
       /citation=[1]
       /citation=[2]
       /evidence=experimental
       2397..2406
       /evidence=experimental
     misc_feature
       1019..1025
       /gene="sgp"
     polyA_signal
       2397..2406
       /evidence=experimental
     ORIGIN
       66.1%; Score 1489.4; DB 14; Length 2406;
       1080
       |||
       1081
       |||
       1082
       |||
       1083
       |||
       1084
       |||
       1085
       |||
       1086
       |||
       1087
       |||
       1088
       |||
       1089
       |||
       1090
       |||
       1091
       |||
       1092
       |||
       1093
       |||
       1094
       |||
       1095
       |||
       1096
       |||
       1097
       |||
       1098
       |||
       1099
       |||
       1100
       |||
       1101
       |||
       1102
       |||
       1103
       |||
       1104
       |||
       1105
       |||
       1106
       |||
       1107
       |||
       1108
       |||
       1109
       |||
       1110
       |||
       1111
       |||
       1112
       |||
       1113
       |||
       1114
       |||
       1115
       |||
       1116
       |||
       1117
       |||
       1118
       |||
       1119
       |||
       1120
       |||
       1121
       |||
       1122
       |||
       1123
       |||
       1124
       |||
       1125
       |||
       1126
       |||
       1127
       |||
       1128
       |||
       1129
       |||
       1130
       |||
       1131
       |||
       1132
       |||
       1133
       |||
       1134
       |||
       1135
       |||
       1136
       |||
       1137
       |||
       1138
       |||
       1139
       |||
       1140
       |||
       1141
       |||
       1142
       |||
       1143
       |||
       1144
       |||
       1145
       |||
       1146
       |||
       1147
       |||
       1148
       |||
       1149
       |||
       1150
       |||
       1151
       |||
       1152
       |||
       1153
       |||
       1154
       |||
       1155
       |||
       1156
       |||
       1157
       |||
       1158
       |||
       1159
       |||
       1160
       |||
       1161
       |||
       1162
       |||
       1163
       |||
       1164
       |||
       1165
       |||
       1166
       |||
       1167
       |||
       1168
       |||
       1169
       |||
       1170
       |||
       1171
       |||
       1172
       |||
       1173
       |||
       1174
       |||
       1175
       |||
       1176
       |||
       1177
       |||
       1178
       |||
       1179
       |||
       1180
       |||
       1181
       |||
       1182
       |||
       1183
       |||
       1184
       |||
       1185
       |||
       1186
       |||
       1187
       |||
       1188
       |||
       1189
       |||
       1190
       |||
       1191
       |||
       1192
       |||
       1193
       |||
       1194
       |||
       1195
       |||
       1196
       |||
       1197
       |||
       1198
       |||
       1199
       |||
       1200
       |||
       1201
       |||
       1202
       |||
       1203
       |||
       1204
       |||
       1205
       |||
       1206
       |||
       1207
       |||
       1208
       |||
       1209
       |||
       1210
       |||
       1211
       |||
       1212
       |||
       1213
       |||
       1214
       |||
       1215
       |||
       1216
       |||
       1217
       |||
       1218
       |||
       1219
       |||
       1220
       |||
       1221
       |||
       1222
       |||
       1223
       |||
       1224
       |||
       1225
       |||
       1226
       |||
       1227
       |||
       1228
       |||
       1229
       |||
       1230
       |||
       1231
       |||
       1232
       |||
       1233
       |||
       1234
       |||
       1235
       |||
       1236
       |||
       1237
       |||
       1238
       |||
       1239
       |||
       1240
       |||
       1241
       |||
       1242
       |||
       1243
       |||
       1244
       |||
       1245
       |||
       1246
       |||
       1247
       |||
       1248
       |||
       1249
       |||
       1250
       |||
       1251
       |||
       1252
       |||
       1253
       |||
       1254
       |||
       1255
       |||
       1256
       |||
       1257
       |||
       1258
       |||
       1259
       |||
       1260
       |||
       1261
       |||
       1262
       |||
       1263
       |||
       1264
       |||
       1265
       |||
       1266
       |||
       1267
       |||
       1268
       |||
       1269
       |||
       1270
       |||
       1271
       |||
       1272
       |||
       1273
       |||
       1274
       |||
       1275
       |||
       1276
       |||
       1277
       |||
       1278
       |||
       1279
       |||
       1280
       |||
       1281
       |||
       1282
       |||
       1283
       |||
       1284
       |||
       1285
       |||
       1286
       |||
       1287
       |||
       1288
       |||
       1289
       |||
       1290
       |||
       1291
       |||
       1292
       |||
       1293
       |||
       1294
       |||
       1295
       |||
       1296
       |||
       1297
       |||
       1298
       |||
       1299
       |||
       1300
       |||
       1301
       |||
       1302
       |||
       1303
       |||
       1304
       |||
       1305
       |||
       1306
       |||
       1307
       |||
       1308
       |||
       1309
       |||
       1310
       |||
       1311
       |||
       1312
       |||
       1313
       |||
       1314
       |||
       1315
       |||
       1316
       |||
       1317
       |||
       1318
       |||
       1319
       |||
       1320
       |||
       1321
       |||
       1322
       |||
       1323
       |||
       1324
       |||
       1325
       |||
       1326
       |||
       1327
       |||
       1328
       |||
       1329
       |||
       1330
       |||
       1331
       |||
       1332
       |||
       1333
       |||
       1334
       |||
       1335
       |||
       1336
       |||
       1337
       |||
       1338
       |||
       1339
       |||
       1340
       |||
       1341
       |||
       1342
       |||
       1343
       |||
       1344
       |||
       1345
       |||
       1346
       |||
       1347
       |||
       1348
       |||
       1349
       |||
       1350
       |||
       1351
       |||
       1352
       |||
       1353
       |||
       1354
       |||
       1355
       |||
       1356
       |||
       1357
       |||
       1358
       |||
       1359
       |||
       1360
       |||
       1361
       |||
       1362
       |||
       1363
       |||
       1364
       |||
       1365
       |||
       1366
       |||
       1367
       |||
       1368
       |||
       1369
       |||
       1370
       |||
       1371
       |||
       1372
       |||
       1373
       |||
       1374
       |||
       1375
       |||
       1376
       |||
       1377
       |||
       1378
       |||
       1379
       |||
       1380
       |||
       1381
       |||
       1382
       |||
       1383
       |||
       1384
       |||
       1385
       |||
       1386
       |||
       1387
       |||
       1388
       |||
       1389
       |||
       1390
       |||
       1391
       |||
       1392
       |||
       1393
       |||
       1394
       |||
       1395
       |||
       1396
       |||
       1397
       |||
       1398
       |||
       1399
       |||
       1400
       |||
       1401
       |||
       1402
       |||
       1403
       |||
       1404
       |||
       1405
       |||
       1406
       |||
       1407
       |||
       1408
       |||
       1409
       |||
       1410
       |||
       1411
       |||
       1412
       |||
       1413
       |||
       1414
       |||
       1415
       |||
       1416
       |||
       1417
       |||
       1418
       |||
       1419
       |||
       1420
       |||
       1421
       |||
       1422
       |||
       1423
       |||
       1424
       |||
       1425
       |||
       1426
       |||
       1427
       |||
       1428
       |||
       1429
       |||
       1430
       |||
       1431
       |||
       1432
       |||
       1433
       |||
       1434
       |||
       1435
       |||
       1436
       |||
       1437
       |||
       1438
       |||
       1439
       |||
       1440
       |||
       1441
       |||
       1442
       |||
       1443
       |||
       1444
       |||
       1445
       |||
       1446
       |||
       1447
       |||
       1448
       |||
       1449
       |||
       1450
       |||
       1451
       |||
       1452
       |||
       1453
       |||
       1454
       |||
       1455
       |||
       1456
       |||
       1457
       |||
       1458
       |||
       1459
       |||
       1460
       |||
       1461
       |||
       1462
       |||
       1463
       |||
       1464
       |||
       1465
       |||
       1466
       |||
       1467
       |||
       1468
       |||
       1469
       |||
       1470
       |||
       1471
       |||
       1472
       |||
       1473
       |||
       1474
       |||
       1475
       |||
       1476
       |||
       1477
       |||
       1478
       |||
       1479
       |||
       1480
       |||
       1481
       |||
       1482
       |||
       1483
       |||
       1484
       |||
       1485
       |||
       1486
       |||
       1487
       |||
       1488
       |||
       1489
       |||
       1490
       |||
       1491
       |||
       1492
       |||
       1493
       |||
       1494
       |||
       1495
       |||
       1496
       |||
       1497
       |||
       1498
       |||
       1499
       |||
       1500
       |||
       1501
       |||
       1502
       |||
       1503
       |||
       1504
       |||
       1505
       |||
       1506
       |||
       1507
       |||
       1508
       |||
       1509
       |||
       1510
       |||
       1511
       |||
       1512
       |||
       1513
       |||
       1514
       |||
       1515
       |||
       1516
       |||
       1517
       |||
       1518
       |||
       1519
       |||
       1520
       |||
       1521
       |||
       1522
       |||
       1523
       |||
       1524
       |||
       1525
       |||
       1526
       |||
       1527
       |||
       1528
       |||
       1529
       |||
       1530
       |||
       1531
       |||
       1532
       |||
       1533
       |||
       1534
       |||
       1535
       |||
       1536
       |||
       1537
       |||
       1538
       |||
       1539
       |||
       1540
       |||
       1541
       |||
       1542
       |||
       1543
       |||
       1544
       |||
       1545
       |||
       1546
       |||
       1547
       |||
       1548
       |||
       1549
       |||
       1550
       |||
       1551
       |||
       1552
       |||
       1553
       |||
       1554
       |||
       1555
       |||
       1556
       |||
       1557
       |||
       1558
       |||
       1559
       |||
       1560
       |||
       1561
       |||
       1562
       |||
       1563
       |||
       1564
       |||
       1565
       |||
       1566
       |||
       1567
       |||
       1568
       |||
       1569
       |||
       1570
       |||
       1571
       |||
       1572
       |||
       1573
       |||
       1574
       |||
       1575
       |||
       1576
       |||
       1577
       |||
       1578
       |||
       1579
       |||
       1580
       |||
       1581
       |||
       1582
       |||
       1583
       |||
       1584
       |||
       1585
       |||
       1586
       |||
       1587
       |||
       1588
       |||
       1589
       |||
       1590
       |||
       1591
       |||
       1592
       |||
       1593
       |||
       1594
       |||
       1595
       |||
       1596
       |||
       1597
       |||
       1598
       |||
       1599
       |||
       1600
       |||
       1601
       |||
       1602
       |||
       1603
       |||
       1604
       |||
       1605
       |||
       1606
       |||
       1607
       |||
       1608
       |||
       1609
       |||
       1610
       |||
       1611
       |||
       1612
       |||
       1613
       |||
       1614
       |||
       1615
       |||
       1616
       |||
       1617
       |||
       1618
       |||
       1619
       |||
       1620
       |||
       1621
       |||
       1622
       |||
       1623
       |||
       1624
       |||
       1625
       |||
       1626
       |||
       1627
       |||
       1628
       |||
       1629
       |||
       1630
       |||
       1631
       |||
       1632
       |||
       1633
       |||
       1634
       |||
       1635
       |||
       1636
       |||
       1637
       |||
       1638
       |||
       1639
       |||
       1640
       |||
       1641
       |||
       1642
       |||
       1643
       |||
       1644
       |||
       1645
       |||
       1646
       |||
       1647
       |||
       1648
       |||
       1649
       |||
       1650
       |||
       1651
       |||
       1652
       |||
       1653
       |||
       1654
       |||
       1655
       |||
       1656
       |||
       1657
       |||
       1658
       |||
       1659
       |||
       1660
       |||
       1661
       |||
       1662
       |||
       1663
       |||
       1664
       |||
       1665
       |||
       1666
       |||
       1667
       |||
       1668
       |||
       1669
       |||
       1670
       |||
       1671
       |||
       1672
       |||
       1673
       |||
       1674
       |||
       1675
       |||
       1676
       |||
       1677
       |||
       1678
       |||
       1679
       |||
       1680
       |||
       1681
       |||
       1682
       |||
       1683
       |||
       1684
       |||
       1685
       |||
       1686
       |||
       1687
       |||
       1688
       |||
       1689
       |||
       1690
       |||
       1691
       |||
       1692
       |||
       1693
       |||
       1694
       |||
       1695
       |||
       1696
       |||
       1697
       |||
       1698
       |||
       1699
       |||
       1700
       |||
       1701
       |||
       1702
       |||
       1703
       |||
       1704
       |||
       1705
       |||
       1706
       |||
       1707
       |||
       1708
       |||
       1709
       |||
       1710
       |||
       1711
       |||
       1712
       |||
       1713
       |||
       1714
       |||
       1715
       |||
       1716
       |||
       1717
       |||
       1718
       |||
       1719
       |||
       1720
       |||
       1721
       |||
       1722
       |||
       1723
       |||
       1724
       |||
       1725
       |||
       1726
       |||
       1727
       |||
       1728
       |||
       1729
       |||
       1730
       |||
       1731
       |||
       1732
       |||
       1733
       |||
       1734
       |||
       1735
       |||
       1736
       |||
       1737
       |||
       1738
       |||
       1739
       |||
       1740
       |||
       1741
       |||
       1742
       |||
       1743
       |||
       1744
       |||
       1745
       |||
       1746
       |||
       1747
       |||
       1748
       |||
       1749
       |||
       1750
       |||
       1751
       |||
       1752
       |||
       1753
       |||
       1754
       |||
       1755
       |||
       1756
       |||
       1757
       |||
       1758
       |||
       1759
       |||
       1760
       |||
       1761
       |||
       1762
       |||
       1763
       |||
       1764
       |||
       1765
       |||
       1766
       |||
       1767
       |||
       1768
       |||
       1769
       |||
       1770
       |||
       1771
       |||
       1772
       |||
       1773
       |||
       1774
       |||
       1775
       |||
       1776
       |||
       1777
       |||
       1778
       |||
       1779
       |||
       1780
       |||
       1781
       |||
       1782
       |||
       1783
       |||
       1784
       |||
       1785
       |||
       1786
       |||
       1787
       |||
       1788
       |||
       1789
       |||
       1790
       |||
       1791
       |||
       1792
       |||
       1793
       |||
       1794
       |||
       1795
       |||
       1796
       |||
       1797
       |||
       1798
       |||
       1799
       |||
       1800
       |||
       1801
       |||
       1802
       |||
       1803
       |||
       1804
       |||
       1805
       |||
       1806
       |||
       1807
       |||
       1808
       |||
       1809
       |||
       1810
       |||
       1811
       |||
       1812
       |||
       1813
       |||
       1814
       |||
       1815
       |||
       1816
       |||
       1817
       |||
       1818
       |||
       1819
       |||
       1820
       |||
       1821
       |||
       1822
       |||
       1823
       |||
       1824
       |||
       1825
       |||
       1826
       |||
       1827
       |||
       1828
       |||
       1829
       |||
       1830
       |||
       1831
       |||
       1832
       |||
       1833
       |||
       1834
       |||
       1835
       |||
       1836
       |||
       1837
       |||
       1838
       |||
       1839
       |||
       1840
       |||
       1841
       |||
       1842
       |||
       1843
       |||
       1844
       |||
       1845
       |||
       1846
       |||
       1847
       |||
       1848
       |||
       1849
       |||
       1850
       |||
       1851
       |||
       1852
       |||
       1853
       |||
       1854
       |||
       1855
       |||
       1856
       |||
       1857
       |||
       1858
       |||
       1859
       |||
       1860
       |||
       1861
       |||
       1862
       |||
       1863
       |||
       1864
       |||
       1865
       |||
       1866
       |||
       1867
       |||
       1868
       |||
       1869
       |||
       1870
       |||
       1871
       |||
       1872
       |||
       1873
       |||
       1874
       |||
       1875
       |||
       1876
       |||
       1877
       |||
       1878
       |||
       1879
       |||
       1880
       |||
       1881
       |||
       1882
       |||
       1883
       |||
       1884
       |||
       1885
       |||
       1886
       |||
       1887
       |||
       1888
       |||
       1889
       |||
       1890
       |||
       1891
       |||
       1892
       |||
       1893
       |||
       1894
       |||
       1895
       |||
       1896
       |||
       1897
       |||
       1898
       |||
       1899
       |||
       1900
       |||
       1901
       |||
       1902
       |||
       1903
       |||
       1904
       |||
       1905
       |||
       1906
       |||
       1907
       |||
       1908
       |||
       1909
       |||
       1910
       |||
       1911
       |||
       1912
       |||
       1913
       |||
       1914
       |||
       1915
       |||
       1916
       |||
       1917
       |||
       1918
       |||
       1919
       |||
       1920
       |||
       1921
       |||
       1922
       |||
       1923
       |||
       1924
       |||
       1925
       |||
       1926
       |||
       1927
       |||
       1928
       |||
       1929
       |||
       1930
       |||
       1931
       |||
       1932
       |||
       1933
       |||
       1934
       |||
       1935
       |||
       1936
       |||
       1937
       |||
       1938
       |||
       1939
       |||
       1940
       |||
       1941
       |||
       1942
       |||
       1943
       |||
       1944
       |||
       1945
       |||
       1946
       |||
       1947
       |||
       1948
       |||
       1949
       |||
       1950
       |||
       1951
       |||
       1952
       |||
       1953
       |||
       1954
       |||
       1955
       |||
       1956
       |||
       1957
       |||
       1958
       |||
       1959
       |||
       1960
       |||
       1961
       |||
       1962
       |||
       1963
       |||
       1964
       |||
       1965
       |||
       1966
       |||
       1967
       |||
       1968
       |||
       1969
       |||
       1970
       |||
       1971
       |||
       1972
       |||
       1973
       |||
       1974
       |||
       1975
       |||
       1976
       |||
       1977
       |||
       1978
       |||
       1979
       |||
       1980
       |||
       1981
       |||
       1982
       |||
       1983
       |||
       1984
       |||
       1985
       |||
       1986
       |||
       1987
       |||
       1988
       |||
       1989
       |||
       1990
       |||
       1991
       |||
       1992
       |||
       1993
       |||
       1994
       |||
       1995
       |||
       1996
       |||
       1997
       |||
       1998
       |||
       1999
       |||
       2000
       |||
       2001
       |||
       2002
       |||
       2003
       |||
       2004
       |||
       2005
       |||
       2006
       |||
       2007
       |||
       2008
       |||
       2009
       |||
       2010
       |||
       2011
       |||
       2012
       |||
       2013
       |||
       2014
       |||
       2015
       |||
       2016
       |||
       2017
       |||
       2018
       |||
       2019
       |||
       2020
       |||
       2021
       |||
       2022
       |||
       2023
       |||
       2024
       |||
       2025
       |||
       2026
       |||
       2027
       |||
       2028
       |||
       2029
       |||
       2030
       |||
       2031
       |||
       2032
       |||
       2033
       |||
       2034
       |||
       2035
       |||
       2036
       |||
       2037
       |||
       2038
       |||
       2039
       |||
       2040
       |||
       2041
       |||
       2042
       |||
       2043
       |||
       2044
       |||
       2045
       |||
       2046
       |||
       2047
       |||
       2048
       |||
       2049
       |||
       2050
       |||
       2051
       |||
       2052
       |||
       2053
       |||
       2054
       |||
       2055
       |||
       2056
       |||
       2057
       |||
       2058
       |||
       2059
       |||
       2060
       |||
       2061
       |||
       2062
       |||
       2063
       |||
       2064
       |||
       2065
       |||
       2066
       |||
       2067
       |||
       2068
       |||
       2069
       |||
       2070
       |||
       2071
       |||
       2072
       |||
       2073
       |||
       2074
       |||
       20
```


Db	1159	ATCATGGCTTCAGAAAATTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1218
Qy	1081	GCAGTGTGCATCTAACAACCCCTTCCCAATCTCCACAGATCCCAATCCCTCACAACC	1140
Db	1219	GCAGTGTGCATCTAACAACCCCTTCCCAATCTCCACAGATCCCAATCCCTCACAACC	1278
Qy	1141	AAACCAAGTTCGGGACAAACAGACCCCATTAATACACCCGTTGATAAATCTTGACATCTCTGAG	1200
Db	1279	AAACCAAGTTCGGGACAAACAGACCCCATTAATACACCCGTTGATAAATCTTGACATCTCTGAG	1338
Qy	1201	GCAACTCAAGTTGAAACAACATCACCGCAGAGACAGACAACAGCAGACAGCCCTCCGACACT	1260
Db	1339	GCAACTCAAGTTGAAACAACATCACCGCAGAGACAGACAACAGCAGACAGCCCTCCGACACT	1398
Qy	1261	CCCTCTGCACACACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC	1320
Db	1399	CCCTCTGCACACACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC	1458
Qy	1321	ACTGACTTCTTGAGCCCGCCACCAACAAGTCCCAAAACCCAGCAGCAGACCGCTGCG	1380
Db	1459	ACTGACTTCTTGAGCCCGCCACCAACAAGTCCCAAAACCCAGCAGCAGACCGCTGCG	1518
Qy	1381	AACAACAACACTCATCACCAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC	1440
Db	1519	AACAACAACACTCATCACCAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC	1578
Qy	1441	TTAATTACCAATATCTTGTGGAGTCGAGACTGATCAGAGCGGAGAGAACTCGA	1500
Db	1579	TTAATTACCAATATCTTGTGGAGTCGAGACTGATCAGAGCGGAGAGAACTCGA	1638
Qy	1501	CGA 1503	
Db	1639	AGA 1641	
RESULT 7			
EVU23187			
LOCUS	EVU23187	2408 bp	RNA linear VRL 28-AUG-2002
DEFINITION	Zaire Ebola virus Mayinga strain glycoprotein (GP) gene, complete cds.		
ACCESSION	U23187		
VERSION	U23187.1	GI:1041204	
KEYWORDS			
SOURCE	Zaire ebolavirus (ZEBOV)		
ORGANISM	Zaire ebolavirus		
REFERENCE	Viruses; serRNA negative-strand viruses; Mononegavirales; Filoviridae; Ebola-like viruses.		
AUTHORS	Sanchez,A., Trappier,S.G., Mahy,B.W., Peters,C.J. and Nichol,S.T.		
TITLE	The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 93 (8), 3602-3607 (1996)		
MEDLINE	96195018		
PUBMED	8622982		
REFERENCE	2 (bases 1 to 2408)		
AUTHORS	Sanchez,A., Trappier,S., Conaty,A.L., Brammer,L., Mahy,B.J.W., Peters,C.J. and Nichol,S.T.		
TITLE	Direct Submision		
JOURNAL	Submitted (22-MAR-1995) Anthony Sanchez, Special Pathogens Branch, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road, Bldg. 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA		
FEATURES	Location/Qualifiers		
source	1. .2408		
	/organism="Zaire ebolavirus"		
	/mol_type="genomic RNA"		
	/strain="Mayinga 1976"		
	/db_xref="taxon:186538"		
	/note="subtype: Zaire"		
	1. .141		
5'UTR	1. .14		
misc_signal	/note="transcription start site"		
gene	142. .2171		

CDS			
	/gene="Gp"		
	join(142. .1027,1027. .2171)		
	/gene="Gp"		
	/functions="cell receptor recognition/binding; structural glycoprotein"		
	/codon_start=1		
	/product="viral spike glycoprotein"		
	/protein_id="AAC54887.1"		
	/db_xref="GI:1041205"		
	/translation="MGVTGILQLPRDRFKRTSPFLMWIIILFORTFSLPLGVHNSTLQ VSDVDKLVCRDKLSSTNQLRSVGLNLENGVATDVPATKRWGPRSGVPKPVNVEAG EWAENCNLEIKKPDGSECLPAAPDGRGPRCRYVHKVSGTGPCAGDPAFHKGAFPL YDRLASTVIYRGTTFAEGVVAFLILPOAKDPFESSHPLREPVNATEDPSSGYSTTI RYQATGFTNETEYLFVLDNLTYVLESREFTPQLQLNETIYTSGRSNTTGLIWK VNPEDITIGEWAFWETKTSLEKFAVKSCLSQLYQTEPKTSVVRVRELLPTQGPQ OLKTTKSLQKIPLOWFKCTKESGLQRI"		
	142. .237		
	/gene="Gp"		
	/functions="unknown"		
	/note="nonstructural glycoprotein; SGP"		
	/codon_start=1		
	/product="small/secreted glycoprotein"		
	/protein_id="AAC54886.1"		
	/db_xref="GI:1041206"		
	/translation="MGVTGILQLPRDRFKRTSPFLMWIIILFORTFSLPLGVHNSTLQ VSDVDKLVCRDKLSSTNQLRSVGLNLENGVATDVPATKRWGPRSGVPKPVNVEAG EWAENCNLEIKKPDGSECLPAAPDGRGPRCRYVHKVSGTGPCAGDPAFHKGAFPL YDRLASTVIYRGTTFAEGVVAFLILPOAKDPFESSHPLREPVNATEDPSSGYSTTI RYQATGFTNETEYLFVLDNLTYVLESREFTPQLQLNETIYTSGRSNTTGLIWK VNPEDITIGEWAFWETKTSLEKFAVKSCLSQLYQTEPKTSVVRVRELLPTQGPQ OLKTTKSLQKIPLOWFKCTKESGLQRI"		
	142. .237		
	/gene="Gp"		
	/note="transcription editing site"		
	2088. .2156		
	/gene="Gp"		
	/note="encodes transmembrane anchor sequence"		
	2398. .2408		
	5'UTR		
	polyA_signal		
	ORIGIN		
	Query Match 66.1%; Score 1489.4; DB 14; Length 2408;		
	Best Local Similarity 99.9%; Pred.No. 0;		
	Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;		
Qy	1	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTACGATTCAGATTCAGAGGACATCATCTTTT	60
Db	142	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTACGATTCAGATTCAGAGGACATCATCTTTT	201
Qy	61	CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT	120
Db	202	CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT	261
Qy	121	AGCACATTACAGGTAGTGATGTCGACAAACTAGTTTGTCTGTGACAACTGTCTATCCACA	180
Db	262	AGCACATTACAGGTAGTGATGTCGACAAACTAGTTTGTCTGTGACAACTGTCTATCCACA	321
Qy	181	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	240
Db	322	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	381
Qy	241	TCTGCAACTAAAGATGGGCTTCAGGTCCGCTTCCCAACAAAGGTGGTCAATATATGAA	300
Db	382	TCTGCAACTAAAGATGGGCTTCAGGTCCGCTTCCCAACAAAGGTGGTCAATATATGAA	441

QY 301 GCTGTGAATGGGCTGAAACCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 442 GCTGTGAATGGGCTGAAACCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG 501
QY 361 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 420
Db 502 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 561
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 480
Db 562 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 621
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGACGACTTTCCTGTAAGGTGTC 540
Db 622 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGACGACTTTCCTGTAAGGTGTC 681
QY 541 GTTGCATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACCACCCCTTGAGA 600
Db 682 GTTGCATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACCACCCCTTGAGA 741
QY 601 GAGCCGCTCAATGCAACGGGAGCCGCTCTAGTGGCTACTATTCTACCAATTAAGATAT 660
Db 742 GAGCCGCTCAATGCAACGGGAGCCGCTCTAGTGGCTACTATTCTACCAATTAAGATAT 801
QY 661 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 802 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 861
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCAGCTGAATGAGACAATA 780
Db 862 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCAGCTGAATGAGACAATA 921
QY 781 TATACAAGTGGGAAAGGAGCAATACCGGGAATACTAATTTGGAAGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAGGAGCAATACCGGGAATACTAATTTGGAAGTCAACCCCGAA 981
QY 841 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAACTAAAAAAACCTCTACTAGAAA 900
Db 982 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAACTAAAAAAACCTCTACTAGAAA 1040
QY 901 ATTGAGTGAAGAGTGTCTTTCAGATTGATCAAGCGAGCCAAAACATCAGTGTG 960
Db 1041 ATTGAGTGAAGAGTGTCTTTCAGATTGATCAAGCGAGCCAAAACATCAGTGTG 1100
QY 961 CAGAGTCCGGCGGCACTTCTTCGACCCAGGAGCAACACACAACTGAAGCACACAAA 1020
Db 1101 CAGAGTCCGGCGGCACTTCTTCGACCCAGGAGCAACACAACTGAAGCACACAAA 1160
QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 1161 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1220
QY 1081 GAGTGTGCGATCTAACAACCTTGGCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
Db 1221 GAGTGTGCGATCTAACAACCTTGGCAATCTCCAGAGTCCCAATCCCTCACAACC 1280
QY 1141 AAACCAAGTCCGAGCAACAGACCCCAATATACACCGTGTATATACTTCACATCTCTGAG 1200
Db 1281 AAACCAAGTCCGAGCAACAGACCCCAATATACACCGTGTATATACTTCACATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAAACAACATCACCGCAGAACAGACAGACAGCCCTCCGACACT 1260
Db 1341 GCAACTCAAGTTGAAACAACATCACCGCAGAACAGACAGACAGCCCTCCGACACT 1400
QY 1261 CCTCTGCAACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC 1320
Db 1401 CCTCTGCAACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC 1460
QY 1321 ACTGACTTCTTGAGCCCGCCACCAACAAGTCCCAAAACACAGCAGCAGCCGTGCG 1380
Db 1461 ACTGACTTCTTGAGCCCGCCACCAACAAGTCCCAAAACACAGCAGCAGCCGTGCG 1520
QY 1381 AACAACAACACTCATCAACAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTTAGGC 1440

Db 1521 AACAAACAACACTCATCAACAGATACCGGAGAAGAGAGTGCACAGCGGGAAGCTAGGC 1580
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCAGGACTGTATCAGCGCGGAGAACTCGA 1500
Db 1581 TTAATTACCAATACTATTGCTGGAGTCCAGGACTGTATCAGCGCGGAGAACTCGA 1640
QY 1501 CGA 1503
Db 1641 AGA 1643
RESULT 8
AF086833 18959 bp RNA linear VRL 28-AUG-2002
LOCUS Zaire Ebola virus strain Mayinga, complete genome.
DEFINITION AF086833
ACCESSION AF086833.2 GI:10141003
VERSION AF086833.2
KEYWORDS Zaire ebolavirus (ZEBOV)
SOURCE Zaire ebolavirus
ORGANISM Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (bases 1 to 18959)
AUTHORS Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses
JOURNAL FEBS Lett. 322 (1), 41-46 (1993)
MEDLINE 93245956
PUBMED 8482365
REFERENCE 2 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Becker,S., Volchkova,V.A., Ternovoj,V.A.,
Kotov,A.N., Netesov,S.V. and Klenk,H.D.
TITLE GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases
JOURNAL Virology 214 (2), 421-430 (1995)
MEDLINE 96130181
PUBMED 8553543
REFERENCE 3 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Feldmann,H., Volchkova,V.A. and Klenk,H.D.
TITLE Processing of the Ebola virus glycoprotein by the proprotein convertase furin
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (10), 5762-5767 (1998)
MEDLINE 98245155
PUBMED 9576958
REFERENCE 4 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Volchkova,V.A., Slenczka,W., Klenk,H.D. and Feldmann,H.
TITLE Release of viral glycoproteins during Ebola virus infection
JOURNAL Virology 245 (1), 110-119 (1998)
MEDLINE 98277077
PUBMED 9614872
REFERENCE 5 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Volchkova,V.A., Chepurinov,A.A., Blinov,V.M., Dolnik,O., Netesov,S.V. and Feldmann,H.
TITLE Characterization of the L gene and 5' trailer region of Ebola virus
JOURNAL J. Gen. Virol. 80 (Pt 2), 355-362 (1999)
MEDLINE 99171703
PUBMED 10073695
REFERENCE 6 (bases 1 to 18959)
AUTHORS Volchkov,V.E.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Institute of Virology, Philipps-University Marburg, Robert-Koch-Str. 17, Marburg 35037, Germany
REFERENCE 7 (bases 1 to 18959)
AUTHORS Volchkov,V.E.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2000) Institute of Virology, Philipps-University Marburg, Robert-Koch-Str. 17, Marburg 35037, Germany
REMARK Sequence update by submitter
COMMENT On Sep 15, 2000 this sequence version replaced gi:4262346.
FEATURES location/Qualifiers
source 1. .18959

```

/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Mayinga"
/db_xref="taxon:186538"
1..55
5'UTR
/citation=[5]
/function="regulation or initiation of RNA replication"
56..3026
/gene="NP"
/gene="NP"
56..3026
/gene="NP"
/product="nucleoprotein"
56..67
/gene="NP"
/translation="putative; transcription start signal"
/citation=[5]
470..2689
/gene="NP"
/function="encapsidation of genomic RNA"
/codon_start=1
/product="nucleoprotein"
/protein_id="AADI4590.1"
/db_xref="GI:4262355"
/translation="WDSPQKIWAPSLTESDMYHKILTAGLSVQOQIVQRVIPVY
QVNLEECQLIQAFEGVDFQESADSFLLMLCLHHAYQGYDKLFESGAVKYLEGH
GRFEVRKRDGKLELLPAVSSGKNIKRTLAAMPHEETTEANAGQFLSPASLFLPK
LVGEGKACLEKVRQIQVHAEQGLIQYPTAQSVGHMMVI FRLMRTNFIKFLLIHQG
MHMVAGHDANDAVISNVAQARFGLIIVKTVLDHILQKTERGVRLHP LARTAKVNE
VNSKALSSLLAKGEYAPFARLLNLGNNLEHGLRPQLSAIAGVATAGHSIQQTV
NVGQYQQLREAYAEKQLOQVAESRELDHLDDQEKILMNFQKNEHSFQQTN
AMVIRKELAKITEAITAASPKTSYHYDDDDIPPPGINDDNPGHQDDDPDSQ
DTTIPDVVDDGSGYQYSYSENGMNA PDLVLDDEDDTKVPNKRSTGGQ
KNSQKIGRQISRTQINVPGHRTIHASAPLTDNDRNFPSGSTSRLMTPTIN
BEAPLDADDESSLPLEDDEQDRDGT SNRTPTVAPVPVYRSEKELPQDE
QQQDQHTOAEARNQDSNTQSEHSPEEMRYHILRSQGPDAVLYTHMMKDBPVFSTSD
CKEYTPDSLEEEYPPWLTEKAMNEENRFTVLQGOQYFVPMHKKKPMALQHQ"
3015..3026
/gene="NP"
3027..3031
/translation="intergenic region"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
/product="VP35"
/citation=[1]
3032..3043
/gene="VP35"
/translation="putative; transcription start signal"
/citation=[1]
3129..4151
/gene="VP35"
/function="polymerase complex protein"
/citation=[1]
/codon_start=1
/product="VP35"
/protein_id="AADI4582.1"
/db_xref="GI:4262347"
/translation="MTTRTKRGHTAANTQDRMPGPESLWSHISEQLMTGRIPVSDIF
CDIENNGLCVYASQCKPKNPKTNSQTQDPCINHSFEEVQVQLASLVQQQTI
ASBSLEQRTISLENGKPYDMAKTISLNRVCAEMVAKYDLLVMTTGRATATAAATE
AYWAHGQPPGPGPSAIGRKIESRDETVPQSVREAFNNLNSSTTSITENFGKPD
ISAKDLRIMYDHLPGFGTAPHOLVQVICKGKDSNLSDI THAEFOASLAEQSDPOCA
LIQITKRVPIQDAAAPPVHIHRSRQDIPRACQKSLRPVPPSPKIDRGWVCVQLQDGK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
/gene="VP40"
/product="VP40"
/citation=[1]
4390..4401
misc_signal

```

```

/gene="VP40"
/translation="transcription start signal"
/citation=[1]
4397..4407
/gene="VP35"
/citation=[1]
4479..5459
/gene="VP40"
/citation=[1]
/codon_start=1
/product="matrix protein"
/protein_id="AADI4583.1"
/db_xref="GI:4262348"
/translation="MRRVILPTAPPEYMEALYVRSNSTIARGNSNTGFLTPEPSVNG
DTFENRPIADDTIDHASHTPGSVSSAFILEAMVNVISGPKVLMKQIPIWPLGVAD
QKTSYFSDSTTAJMLASYTTTHFGKATNPLVRNRLGFGIPDHLRLLRIGNQAPLQLE
FVLPPVQYFTFDLTALKLIITQPLPAATWDDTPTGSGALRPGISFHKLRPILL
PNKSGKGNAGDLTSPKIQAIMTSLQDKVIPDPTKNIMGIVBPTLVHKLTKKKV
TSKNGQPIIPVLLPKYIGLDPVAPGDLTMTITQDCDTCCHSPASLPAVIEK"
5883..5894
/gene="VP40"
/citation=[1]
5895..5899
/translation="intergenic region"
5900..8305
/gene="GP"
5900..8305
/gene="GP"
/product="sGP"
/translation="unedited mRNA"
/citation=[2]
5900..5911
/gene="GP"
/translation="putative; transcription start signal"
/citation=[2]
join(6039..6923,6923..8068)
/gene="GP"
/function="receptor binding and fusion"
/translation="an addition A residue is inserted during
transcription; encodes two disulfide linked subunits GP1
and GP2"
/citation=[2]
/citation=[3]
/citation=[4]
/codon_start=1
/exception="RNA editing"
/product="viral spike glycoprotein precursor"
/protein_id="AADI4585.1"
/db_xref="GI:4262350"
/translation="MGVTGILQLPRDKRTSPFLMWIILFQRTFSPLGVHNSTLQ
VSDVDKLVCRDKLSSTNQLRSLGNLGNVATDVPSATKRWGFRSPVPPVNVYEAQ
EWAECYNLEIKKPDGSECLPAAPDGIKRGPRYVHKVSGTGCAGDFAFHKGAF
LYDLASTVLYRGTTFAGVVAFLILPAQKDPFSSHPLREPVNATDPSSGYSTTI
RYQATGTEVETELFEVDNLTVQLSRTPTQFLLOLNETIYSGKSNSTGKLIWK
VNEPDITIGWAFWETKNTLRKIRSELSFTVVSNGAKNISQSPARTSSDPGNT
TTEDHKINAGSNAMQVHSQGREAAVSHLTITATISTSPQSTTKPGPNSTHTPT
VYKLDISEATQVEQHEHRTNDSTPSATTAAGPPKAEKNTNTSKSTDLDPAATT
SPONSHETAGNNTHQDTGEESSAGKGLITWTIAGVAGLITGRRTRREALVNAQ
PKCNPLHWTQDEGAIGLAWIPYFGPAEGVYIEGLMNDQLICGLRLQANETT
QALQFLRATTELRTFSILNRKAIDFLLORWGTCTHILGPCCCTCPHDTWTKNTDKID
QIILHDFVDKTLFDQGDNDNDNMWTRQWIPAGIGVTGVIIAIVLFCICKFVF"
7529..7540
/gene="GP"
/translation="encodes the glycoprotein cleavage site, precursor
GP is cleaved by subtilisin-like cellular protease furin
into subunits GP1 and GP2 that are linked by a disulfide
bond"
/citation=[3]

```

```

Query Match      66.1%; Score 1489.4; DB 14; Length 18959;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

Qy	1	ATGGGCGTTACAGGAATATTGCAAGTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT	60
Db	6039	ATGGGCGTTACAGGAATATTGCAAGTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT	6098
Qy	61	CTTTGGGTAAATTATCTCTTTTCCAAGAAACATTTTCCATCCACCTTGGAGTCATCCACAAT	120
Db	6099	CTTTGGGTAAATTATCTCTTTTCCAAGAAACATTTTCCATCCACCTTGGAGTCATCCACAAT	6158
Qy	121	AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTGCATCCACA	180
Db	6159	AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTGCATCCACA	6218
Qy	181	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGGCCA	240
Db	6219	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGGCCA	6278
Qy	241	TCTGCAACTAAAGATCGGGCTTCAGGTGCGGTGCCACCAAGGTGGTCAATTATGAA	300
Db	6279	TCTGCAACTAAAGATCGGGCTTCAGGTGCGGTGCCACCAAGGTGGTCAATTATGAA	6338
Qy	301	GCTGGTGAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAAACCTGACGGAGTGAG	360
Db	6339	GCTGGTGAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAAACCTGACGGAGTGAG	6398
Qy	361	TGTTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA	420
Db	6399	TGTTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA	6458
Qy	421	GTATCAGGACGGGACCGTGTGCCGGAGACTTTGGCCTTCCATAAAGAGGGTGCTTTCTTC	480
Db	6459	GTATCAGGACGGGACCGTGTGCCGGAGACTTTGGCCTTCCATAAAGAGGGTGCTTTCTTC	6518
Qy	481	CTGTATGATCGACTTGCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTCGAAGGTGTC	540
Db	6519	CTGTATGATCGACTTGCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTCGAAGGTGTC	6578
Qy	541	GTTCGATTTCTGATPACTGCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA	600
Db	6579	GTTCGATTTCTGATPACTGCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA	6638
Qy	601	GAGCCGTCATGCAACGGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT	660
Db	6639	GAGCCGTCATGCAACGGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT	6698
Qy	661	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC	720
Db	6699	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC	6758
Qy	721	TACGTCCAACCTTGAATCAAGATTCCACCACAGTTCTGCTCCAGCTGGAATGAGACAATA	780
Db	6759	TACGTCCAACCTTGAATCAAGATTCCACCACAGTTCTGCTCCAGCTGGAATGAGACAATA	6818
Qy	781	TATACAGTGGGAAAAGAGCAATACACCGGAAAACTAATTTGGAAAGGTCAACCCCGAA	840
Db	6819	TATACAGTGGGAAAAGAGCAATACACCGGAAAACTAATTTGGAAAGGTCAACCCCGAA	6878
Qy	841	ATTGATACAACATCGGGAGTGGGCCCTCTGGGAAACTAAAAAACCCTCACTAGAAAA	900
Db	6879	ATTGATACAACATCGGGAGTGGGCCCTCTGGGAAACT-AAAAAACCCTCACTAGAAAA	6937
Qy	901	ATTCGCAAGTGAAGATTGTCTTTTCCACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT	960
Db	6938	ATTCGCAAGTGAAGATTGTCTTTTCCACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT	6997
Qy	961	CAGAGTCGGCGGAACTTTCTTCGACCCGAGGCCAACACAACTGAAGACCAAAA	1020
Db	6998	CAGAGTCGGCGGAACTTTCTTCGACCCGAGGCCAACACAACTGAAGACCAAAA	7057
Qy	1021	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT	1080
Db	7058	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT	7117
Qy	1081	GCAGTGTGCAATCTAAACAACCCCTTGGCCACAATCTCCACGAGTCCCAATGCCTCAACAAC	1140

Db	7118	GCAGTGTGCGCATCTAAACAACCCCTTGCCACAATCTCCACGAGTCCCCAAATCCCTTCAACAACC	71177
Qy	1141	AAACACAGGTCGGACAAACAGCACCACATAATACACCCGCTGTATAAACTTGACATCTCTGAG	1200
Db	7178	AAACACAGGTCGGACAAACAGCACCACATAATACACCCGCTGTATAAACTTGACATCTCTGAG	7237
Qy	1201	GCAACTCAAGTTGAAACAAATACACCGAGAACAGACAAACAGACAGCAGCCTCCGACACT	1260
Db	7238	GCAACTCAAGTTGAAACAAATACACCGAGAACAGACAAACAGACAGCAGCCTCCGACACT	7297
Qy	1261	CCCTCTGCCACGACCGCAGCCGAGCCGACCAACAAAGTCCCAAAACCAACGACGACGCAAGAGC	1320
Db	7298	CCCTCTGCCACGACCGCAGCCGAGCCGACCAACAAAGTCCCAAAACCAACGACGACGCAAGAGC	7357
Qy	1321	ACTGACTTCTCGGACCCCGCCACCAACAAAGTCCCAAAACCAACGACGACGACGCTGGC	1380
Db	7358	ACTGACTTCTCGGACCCCGCCACCAACAAAGTCCCAAAACCAACGACGACGACGCTGGC	7417
Qy	1381	AACAACAACACTCATCAACCAAGTACCGGAGAGAGAGTGCACGACGCGGAGAGCTAGGC	1440
Db	7418	AACAACAACACTCATCAACCAAGTACCGGAGAGAGAGTGCACGACGCGGAGAGCTAGGC	7477
Qy	1441	TTAATTACCAATCTATTGCTGAGTGCAGGACTCATACAGGCGGAGAGAACTCGA	1500
Db	7478	TTAATTACCAATCTATTGCTGAGTGCAGGACTCATACAGGCGGAGAGAACTCGA	7537
Qy	1501	CGA 1503	
Db	7538	AGA 7540	
RESULT 9			
AF272001			
LOCUS	AF272001	19959 bp	RNA linear VRL 28-AUG-2002
DEFINITION	Zaire Ebola virus strain Mayinga, complete genome.		
ACCESSION	AF272001		
VERSION	AF272001.1	GI:11761745	
KEYWORDS			
SOURCE	Zaire ebolavirus (ZEBOV)		
ORGANISM	Zaire ebolavirus		
REFERENCE	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Ebola-like viruses.		
AUTHORS	1 (bases 1 to 18959) Volchkov, V.E., Volchkova, V.A., Chepurinov, A.A., Blinov, V.M., Dolnik, O., Netesov, S.V. and Feldmann, H.		
TITLE	Characterization of the L gene and 5' trailer region of Ebola virus		
JOURNAL	J. Gen. Virol. 80 (Pt 2), 355-362 (1999)		
MEDLINE	99171703		
PUBMED	10073695		
REFERENCE	2 (bases 1 to 18959) Volchkov, V.E., Chepurinov, A.A., Volchkova, V.A., Ternovoj, V.A. and Klenk, H.D.		
AUTHORS	Molecular characterization of guinea pig-adapted variants of Ebola virus		
TITLE	Virology 277 (1), 147-155 (2000)		
JOURNAL	20517553		
MEDLINE	11062045		
PUBMED	3 (bases 1 to 18959) Volchkov, V.E.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (24-MAY-2000) Institute of Virology, Robert-Koch-Str.17, Marburg 35037, Germany		
JOURNAL	Location/Qualifiers		
FEATURES	1. .18959 /organism="Zaire ebolavirus" /viroion /mol_type="genomic RNA" /strain="Mayinga" /db_xref="taxon:186538" /notes="subtype: Zaire; biotype: Guinea pig adapted variant 8mc"		
source	1. .55		
misc feature	1. .55		


```
misc_feature      8288..8299
/gene="VP30"
/feature="transcription start signal"
polyA_signal      8295..8305
/gene="VP30"
CDS               8509..9375
/gene="VP30"
/note="nucleoprotein"
/codon_start=1
/product="VP30"
/protein_id="AAG40169.1"
/db_xref="GI:11761751"
/translation="MEASLYERPRRAARQSHRSDHHDHVRASSRRENYRGEYRQSR
ASQVRPTVFHKRRVPLTPPPAPKDIPTLKLKGLCDSSFCCKKHQLESITDRLLL
LIARNTCSVEQLNITAPKDSRLANPTADDDPQEEGPKITLLTLIKTAHWARQDIR

Query Match      66.1%; Score 1489.4; DB 14; Length 18959;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTTTT 60
DB |
QY 6039 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTTTT 6098
DB |
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 120
DB |
QY 6099 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 6158
DB |
QY 121 AGCACATTTACAGGTAGTGAATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
DB |
QY 6159 AGCACATTTACAGGTAGTGAATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 6218
DB |
QY 181 AATCAATTCAGATCAGTTCGATCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 240
DB |
QY 6219 AATCAATTCAGATCAGTTCGATCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 6278
DB |
QY 241 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 300
DB |
QY 6279 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 6338
DB |
QY 301 GCTGTGAATGGCTGAAATCTGCTACAACTCTTGAATCAAAAACCTGACGGGAGTGAG 360
DB |
QY 6339 GCTGTGAATGGCTGAAATCTGCTACAACTCTTGAATCAAAAACCTGACGGGAGTGAG 6398
DB |
QY 361 TGTCTACAGACGCGCAGACGGGATTCGGGCTTCCCGCGTTCGGGTATGTGCACAAA 420
DB |
QY 6399 TGTCTACAGACGCGCAGACGGGATTCGGGCTTCCCGCGTTCGGGTATGTGCACAAA 6458
DB |
QY 421 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480
DB |
QY 6459 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 6518
DB |
QY 481 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 540
DB |
QY 6519 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 6578
DB |
QY 541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTCAGCTTCACACCCCTTGAGA 600
DB |
QY 6579 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTCAGCTTCACACCCCTTGAGA 6638
DB |
QY 601 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCACAATATAGATAT 660
DB |
QY 6639 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCACAATATAGATAT 6698
DB |
QY 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTCTGAGTTGACAAATTTGACC 720
DB |
QY 6699 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTCTGAGTTGACAAATTTGACC 6758
DB |
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACCAAGTTTCTGCTCAGCTGAATGAGACAATA 780
DB |
QY 6759 TAGCTCCAACTTGAATCAAGATTTCACACCAAGTTTCTGCTCAGCTGAATGAGACAATA 6818
DB |
QY 781 TATACAAGTGGGAAAAGGAGCAATACCAACGGGAAAATAATTTGGGAAGGTCAACCCCGAA 840
```

```
DB 6819 TATACAAGTGGGAAAAGGAGCAATACCAACGGGAAAATAATTTTGGGAAGGTCAACCCCGAA 6878
QY 841 ATTGATACAAATCGGGAGTGGGCTTCTTGGGAAACTTAAATAAAAACTCTACTAGAAAA 900
DB 6879 ATTGATACAAATCGGGAGTGGGCTTCTTGGGAAACTTAAATAAAAACTCTACTAGAAAA 6937
QY 901 ATTCGAGTGAAGAGTTTCTTTTTCACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGT 960
DB 6938 ATTCGAGTGAAGAGTTTCTTTTTCACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGT 6997
QY 961 CAGAGTCCGGCGCGAACTTCTTTCGACCCAGGAGCAACAACTGAGAGACACAAAA 1020
DB 6998 CAGAGTCCGGCGCGAACTTCTTTCGACCCAGGAGCAACAACTGAGAGACACAAAA 7057
QY 1021 ATCATGGCTTCAGAAAAATCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB 7058 ATCATGGCTTCAGAAAAATCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAAGCT 7117
QY 1081 GCAGTGTGCGCATCTAAACAACCTTTGCCACAATCTCCACGAGTCCCAATCCCTCACAACC 1140
DB 7118 GCAGTGTGCGCATCTAAACAACCTTTGCCACAATCTCCACGAGTCCCAATCCCTCACAACC 7177
QY 1141 AAACAGGTCCGGAACAACAGACCCATTAATACACCCGTGTATAAACTTGACATCTCTGAG 1200
DB 7178 AAACAGGTCCGGAACAACAGACCCATTAATACACCCGTGTATAAACTTGACATCTCTGAG 7237
QY 1201 GCACCTCAAGTTGAACAACATCACCGCAGAACAGACAACAGCAGCAGCCTCCGACACT 1260
DB 7238 GCACCTCAAGTTGAACAACATCACCGCAGAACAGACAACAGCAGCAGCCTCCGACACT 7297
QY 1261 CCCTCTGCGACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAGAGC 1320
DB 7298 CCCTCTGCGACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAGAGC 7357
QY 1321 ACTGACTTCTGAGACCCCGCACCAACAAGTCCCAAAACCAACGAGCAGAGCAGCCTGSC 1380
DB 7358 ACTGACTTCTGAGACCCCGCACCAACAAGTCCCAAAACCAACGAGCAGAGCAGCCTGSC 7417
QY 1381 AACACAACTCTATCAACAAGATACCGAGAGAGAGTCCAGCAGCGGGAAGCTTAGGC 1440
DB 7418 AACACAACTCTATCAACAAGATACCGAGAGAGAGTCCAGCAGCGGGAAGCTTAGGC 7477
QY 1441 TTAATTACCAATATCTATTTGCTGGAGTCCAGGACTGATCAGAGCGGGGAGAACTCGA 1500
DB 7478 TTAATTACCAATATCTATTTGCTGGAGTCCAGGACTGATCAGAGCGGGGAGAACTCGA 7537
QY 1501 CGA 1503
DB 7538 AGA 7540

RESULT 10
AY142960
LOCUS      18959 bp      RNA linear      VRL 16-OCT-2002
DEFINITION Zaire Ebola virus strain Mayinga subtype Zaire, complete genome.
ACCESSION  AY142960
VERSION     AY142960.1  GI:23630482
KEYWORDS
SOURCE      Zaire ebolavirus (ZEBOV)
ORGANISM    Zaire ebolavirus
            Viruses; ssRNA negative-strand viruses; Mononegavirales;
            Filoviridae; Ebola-like viruses.
REFERENCE   1 (bases 1 to 18959)
AUTHORS     Wilson,J.A., Koudig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE       Ebola virus subtype Zaire strain Mayinga complete genome
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 18959)
AUTHORS     Wilson,J.A., Koudig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE       Direct Submission
JOURNAL     Submitted (16-AUG-2002) Virology, USAMRIID, 1425 Porter St.,
            Frederick, MD 21702, USA
FEATURES
            Location/Qualifiers
```


AAIDFSRTPSGKRRLQILGYLEGRTRTLASKIINNNTETPVLDRLAKITLQRWSLWFSY
LBDNIIAALATQITCTVDLAQILREYSWAHLEGRPLIGATLPCMEQKRFVWLKP
YBQCQCNSKQPKFPVSAVKKHIVSWAPNAGRIWITGDGPIYIGSRDXTKIGQ
PAIKPCPSAALREALELASRLTWTQGSNSDLIKPFLAARNVLSVQELIOMTPSH
YSGNIVHRYNQTSYSPHSFANRMSNATRLVSTNLTGFEFGGGSQSDSNLIIFQNV
NTAVAFDILKFRNTSATDIOYNRAHLHLTKCCTREVPAQILYITSTLDDLUYKRENE
LIYDNPLKGGNLNCSIFDNPPFQKRLNIIEDDLIRPLHSGWELAKTIMQSIISDS
NNSDPTSSGSTRFTFLTPYKIGLLYSFAFVSYLVNLTIRTKLTLNDPLFY
LTTQIHNLPRLSLIKLPFKHASVMSRLMSIDPHFSYVIGGAAGDGLSDAARLPLR
TSISPLTPKEMWINRGITVPLWIVYPLEGONPTPVNNFLYQIVELAVHDSRROQAF
KTTISDHVPHDNLVYTCCKSTASNPFLASLAWRSRHRNSNRKYLARSSSTGSSNNNS
DNHIERQSTTRPDHGDTRNLVLRQHEIKRTTIPQENTHQGFSQFSLDSACGT
ANPKLFORSNRVDFQDHSNASKRGHQIISHLRVLFPFLLSQSTROUTSNESQTO
DEISYKOLRSLVITTYVCRFTGIVSSMHYKLDVLEWELIESFKNADTLAEGEGAL
LLIKYQVLTFLNTLATYESSIEISVSMGTMRLLPMSKFNHQDIIEIILNNSAQ
LTDITNPTWFKDQRLPKQEVITMDAETTENINRSKLYEAVYKLIHLIDPVSVLKA
VVLKFLSDTEQMLNLDNLPFPATGYLTKPITSSARSSEWYLCITNLFSTTRKMPH
QNHLSCKVILTALQIQRSFYWLSHLTYQVADCSLHUSYIKRGFFSLEKYLHRYNL
VDSKRGPLVSIITQHLAHLRAEIRELTNDYNQORSTQTYHPRTAKGRITKLVDNYL
KFLIVQALKHNGTWOAEFKLPPELISVGNRFYHIRDNCBERFLVQTLYLHMRQDSE
VKLIBRLTGLLSLFPDGLYRFD"

ORIGIN

Query Match 66.1%; Score 1489.4; DB 14; Length 18959;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	1	ATGGGGTTTACAGGATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTCTTT	60
DB	6039	ATGGGGTTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTCTTT	6098
QY	61	CTTTGGGTAATTATCTCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT	120
DB	6099	CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT	6158
QY	121	AGCACATTAACAGTTAGTGTGACAAAATAGTTTGTGCGTGAACAACTGTCATCCACA	180
DB	6159	AGCACATTAACAGTTAGTGTGACAAAATAGTTTGTGCGTGAACAACTGTCATCCACA	6218
QY	181	AATCAATTCAGATCAGTTGCGATGTAATCTCGAAGGGAATGGAGTGCACCTGCGCA	240
DB	6219	AATCAATTCAGATCAGTTGCGATGTAATCTCGAAGGGAATGGAGTGCACCTGCGCA	6278
QY	241	TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGCTCAATTATGAA	300
DB	6279	TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGCTCAATTATGAA	6338
QY	301	GCTGTGAAATGGGCTGAAAATCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG	360
DB	6339	GCTGTGAAATGGGCTGAAAATCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG	6398
QY	361	TGTTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCGGTATCTGTCACAAA	420
DB	6399	TGTTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCGGTATCTGTCACAAA	6458
QY	421	GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCGTTCATATAAGAGGGTCTTTCTTC	480
DB	6459	GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCGTTCATATAAGAGGGTCTTTCTTC	6518
QY	481	CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC	540
DB	6519	CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC	6578
QY	541	GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTTCACACCCCTTGAGA	600
DB	6579	GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTTCACACCCCTTGAGA	6638
QY	601	GAGCCGGTCAATGCAACCGGAGACCGCTGTAGTGGCTACTATTCTACCAACAATTAGATAT	660
DB	6639	GAGCCGGTCAATGCAACCGGAGACCGCTGTAGTGGCTACTATTCTACCAACAATTAGATAT	6698
QY	661	CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCAGGTTGACAAATTTGACC	720

DB	6699	CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC	6758
QY	721	TACGTCCAACTTGAATCAAGATTTCACACACAGATTTCCTGCTCCAGCTGAATGAGACAATA	780
DB	6759	TACGTCCAACTTGAATCAAGATTTCACACACAGATTTCCTGCTCCAGCTGAATGAGACAATA	6818
QY	781	TATACAAGTGGGAAAGAGGACAAATACCAACGGGAAACATAATTTGGAAAGTCAACCCCGAA	840
DB	6819	TATACAAGTGGGAAAGAGGACAAATACCAACGGGAAACATAATTTGGAAAGTCAACCCCGAA	6878
QY	841	ATTGATACAAACAAATCGGGAGTGGGCTTCTGGGAAACCTAAAAAATCCTACTAGAAAA	900
DB	6879	ATTGATACAAACAAATCGGGAGTGGGCTTCTGGGAAACCT-AAAAAATCCTACTAGAAAA	6937
QY	901	ATTCGAGTGAAGAGTTCCTTTTTCACAGTTGTATCAAAACGGAGCCAAAAAATCATCTAGTGT	960
DB	6938	ATTCGAGTGAAGAGTTCCTTTTTCACAGTTGTATCAAAACGGAGCCAAAAAATCATCTAGTGT	6997
QY	961	CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAAACAACTGAAGACACACAAA	1020
DB	6998	CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAAACAACTGAAGACACACAAA	7057
QY	1021	ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTTCAAGTGCACAGTCAAGGAGGGAAGCT	1080
DB	7058	ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTTCAAGTGCACAGTCAAGGAGGGAAGCT	7117
QY	1081	GCAGTGTGCGCATCTAAACAACTTTCGCCACAATCTCCAGAGTCCCAATTCCTCAACACC	1140
DB	7118	GCAGTGTGCGCATCTAAACAACTTTCGCCACAATCTCCAGAGTCCCAATTCCTCAACACC	7177
QY	1141	AAACCAGTTCGGGACCAACAGCACCCCATTAATACACCGTGTATAAATCTTGACATCTCTGAG	1200
DB	7178	AAACCAGTTCGGGACCAACAGCACCCCATTAATACACCGTGTATAAATCTTGACATCTCTGAG	7237
QY	1201	GCACTCAAGTTGAACACATCAACGAGAACAGACAGACAGACAGACCTCCGACACT	1260
DB	7238	GCACTCAAGTTGAACACATCAACGAGAACAGACAGACAGACAGACCTCCGACACT	7297
QY	1261	CCCTCTGCACGACCGGACCGGACCCCAAAAGCAGAGAACCAACCAACGAGCAAGAGC	1320
DB	7298	CCCTCTGCACGACCGGACCGGACCCCAAAAGCAGAGAACCAACCAACGAGCAAGAGC	7357
QY	1321	ACTGATTTCTGGACCCCGCCCAACCAACAAAGTCCCAAAACCAACGAGGAGACCGTGGC	1380
DB	7358	ACTGATTTCTGGACCCCGCCCAACCAACAAAGTCCCAAAACCAACGAGGAGACCGTGGC	7417
QY	1381	PACACACACTCATCACACAGATACCGGAGAGAGTGCACGAGCGGGAGGAGTGGC	1440
DB	7418	PACACACACTCATCACACAGATACCGGAGAGAGTGCACGAGCGGGAGGAGTGGC	7477
QY	1441	TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA	1500
DB	7478	TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA	7537
QY	1501	CGA 1503	
DB	7538	AGA 7540	

RESULT 11

EVU81161

LOCUS

DEFINITION

Zaire Ebola virus

non-structural glycoprotein sgp (GP) gene, complete cds.

ACCESSION

U81161

VERSION

U81161.1

KEYWORDS

ORGANISM

Zaire ebolavirus (ZEBOV)

SOURCE

Viruses; ssRNA negative-strand viruses; Mononegavirales;

Filoviridae; Ebola-like viruses.

REFERENCE

1 (scies)

AUTHORS

Volchikov, V. E., Blinov, V. M., Kotov, A., Cherpurnov, A. and Netesov, S.

The full-length nucleotide sequence of the Ebola virus
 Thesis (1993) Institut f. Virologie, Philipps-Universitaet
 2 (sites)
 AUTHORS
 Volchkov, V., Volchkova, V., Eckel, C., Klenk, H. D., Bouloy, M.,
 LeGuennou, B. and Feldmann, H.
 Emergence of subtype Zaire Ebola virus in Gabon
 Virology 232 (1), 139-144 (1997)
 97329084
 MEDLINE
 PUBMED
 REFERENCE
 3 (bases 1 to 2408)
 Volchkov, V. E.
 Direct Submission
 Submitted (06-DEC-1996) Volchkov V. E., Philipps-Universitaet,
 Institut f. Virologie, Robert-Koch-Str.17, Marburg, Germany, 35037
 FEATURES
 Location/Qualifiers
 1. .2408
 /organism="Zaire ebolavirus"
 /mol_type="genomic RNA"
 /strain="Eckron-76"
 /db_xref="taxon:186538"
 3. .13
 /function="transcriptional start signal for GP mRNA"
 142. .2171
 /gene="GP"
 142. .2171
 /gene="GP"
 /function="structural protein"
 /function="cell receptor recognition"
 /function="fusion"
 /notes="full-length glycoprotein synthesized from edited
 mRNA species; an additional A residue is inserted at the
 editing site during GP gene transcription"
 /codon_start=1
 /evidence=experimental
 /exception="RNA editing"
 /product="virion spike glycoprotein"
 /protein_id="AAC57992.1"
 /db_xref="GI:1753171"
 /translation="MGVTGLQLPRDRFKRTSPFLWILFQRTFSPILGVHNSTLQ
 VNDVNDKLRDKLSNQLRSVGLNLEGVATDVPATKRMGRGSPGVPKVNVEAG
 EWAENKYLEIKKPDGSECLPAIDGIRGPPRCYVHKVSGTGPCAGDPAFKHEGAF
 LYDLASTVIYRGTTAEGVVAFLIDPQAKDFSSHLPREPVNATEDPSSGYSTITI
 RYQATGFTGNETEYLFEVDNLTYVQLSERTFPQFLQLNETIYTSKRSNTTCKLIWK
 VNPEDITIGTEWAFWTKNLRKIRSELSFTVNSGAKNISQGPARTSSDPGNT
 TTEDHKIMASENSAMVQSHQREAAVSHLATTATISTSPQSLTTPKPDNDSTHTPT
 KYKLDISEATQVEQHRRTDNTASDTPSATTAAGPKAENTNTSKSTDFELDPATT
 SQNHSETAGNNTHHQDTGESASGSKGLINTIAGVAGLITGRRTRREAINNAQ
 PCPNPLNHYTQDEGAALGLAWIYFPAAGIYTEGLHMNQNLGICLRQLANET
 QALQLFRATTELRTSLILNRKAIDFLQRWGSGTCHILGPDCCIEPHDWTQNTDKIK
 QIITHFDVKTLPDQGDNDNNWTTGWRQWIPAGIVGTGVIIAIVAFICIKFVF"
 142. .1236
 /gene="GP"
 /notes="synthesized from non-edited GP mRNA species."
 /codon_start=1
 /evidence=experimental
 /product="small secreted non-structural glycoprotein sGP"
 /protein_id="AAC57993.1"
 /db_xref="GI:1753172"
 /translation="MGVTGLQLPRDRFKRTSPFLWILFQRTFSPILGVHNSTLQ
 VNDVNDKLRDKLSNQLRSVGLNLEGVATDVPATKRMGRGSPGVPKVNVEAG
 EWAENKYLEIKKPDGSECLPAIDGIRGPPRCYVHKVSGTGPCAGDPAFKHEGAF
 LYDLASTVIYRGTTAEGVVAFLIDPQAKDFSSHLPREPVNATEDPSSGYSTITI
 RYQATGFTGNETEYLFEVDNLTYVQLSERTFPQFLQLNETIYTSKRSNTTCKLIWK
 VNPEDITIGTEWAFWTKNLSLEKPAVSKLSLYQTEPKTSVVRVRELLPTQGPSTQ
 QLTKTSWQKQIKPLQWFKCTVKEGKLCQRI"
 2391. .2402
 /function="transcriptional start signal for VP30 mRNA"
 2398. .2408
 misc_signal
 polyA_signal
 ORIGIN
 Query Match 66.1%; Score 1487.8; DB 14; Length 2408;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1500; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 1081 GCAGTGTGCGCATCTAACCAACCCCTTGCCCAACATCTCCACGAGTCCCAATCCCTCAACCC 1140
Db 1221 GCAGTGTGCGCATCTAACCAACCCCTTGCCCAACATCTCCACGAGTCCCAATCCCTCAACCC 1280
QY 1141 AAACAGGTCCGGACACAGCACCCATATACACCGGTGTATATAACTTGACATCTCTGAG 1200
Db 1281 AAACAGGTCCGGACACAGCACCCATATACACCGGTGTATATAACTTGACATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAAACAGACAAACAGACAGACAGCCTCCGACACT 1260
Db 1341 GCAACTCAAGTTGAACAACATCACCGCAGAAACAGACAAACAGACAGACAGCCTCCGACACT 1400
QY 1261 CCCTCTGCCACACGCGCAGCGGACCCCAAGCAGAGAAACACCAACAGCAGCAGAGC 1320
Db 1401 CCCTCTGCCACACGCGCAGCGGACCCCAAGCAGAGAAACACCAACAGCAGCAGAGC 1460
QY 1321 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCGCCAAAACCAACAGCAGCAGCCTGGC 1380
Db 1461 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCGCCAAAACCAACAGCAGCAGCCTGGC 1520
QY 1381 AACAAACACATCATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGAAGCTAGGC 1440
Db 1521 AACAAACACATCATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGAAGCTAGGC 1580
QY 1441 TTAATTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA 1500
Db 1581 TTAATTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA 1640
QY 1501 CGA 1503
Db 1641 AGA 1643

RESULT 12
AF499101
LOCUS AF499101 18960 bp RNA linear VRL 28-AUG-2002
DEFINITION Zaire Ebola virus strain Mayinga, complete genome.
ACCESSION AF499101
VERSION AF499101.1 GI:21702647
KEYWORDS Zaire ebolavirus (ZEBOV)
SOURCE Zaire ebolavirus
ORGANISM Viruses; sRNA negative-strand viruses; Mononegavirales;
REFERENCE 1 (bases 1 to 18960)
AUTHORS Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE Mouse adapted variant of Ebola virus subtype Zaire strain Mayinga
complete genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 18960)
AUTHORS Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Diagnostic System Division, USAMRIID, 1425
Porter St., Ft. Detrick, Frederick, MD 21702, USA
FEATURES
source
1..18960
/organism="Zaire ebolavirus"
/virion
/mol_type="genomic RNA"
/db_xref="taxon:186538"
/strain="Mayinga"
/notes="subtype: Zaire; biotype: mouse adapted variant"
misc_feature 1..55
/notes="putative leader region"
gene 56..3026
/gene="NP"
mRNA 56..3026
/gene="NP"
/misc_signal 56..67
/notes="nucleoprotein NP"
misc_signal 56..67
/notes="transcription start signal"
CDS 470..2689
```

```
/gene="NP"
/codon_start=1
/product="nucleoprotein NP"
/protein_id="AAW76031.1"
/db_xref="GI:21702648"
/translation="MDSRPQKIWMAPSLTESDMYHKILKTAGLSVQQGIIVRQVITPVY
QVNNLEICQLIIQAFAGVDVQBSADGFLMLCLHHAYQGDYKLFLESGAVKYLEGH
GFRFEVKRGRGKLELLPAVSSGKNIKRTLAAMPEBETTEANAGQFLSFASLFLPK
LVGEKACLEKVRQIOVHAEOGLIOYPTAQSVGHMMVI FRLMRTNFIKFLILHQG
MHVYAGHDANDAVTNSVAOARFSGLLIVKTLHILLOKTBGVRLHPLARTAKVKNIE
VNSFKALSSLAKEGYAPFARLLNLSGVNLEHGLFQLSAIALGVATAGSTIAGV
NVSGYQQLREATAEAEKQLQYAEARELDHLGDDDEKKILMNFHQKNEISFQTN
AMVTLRKERLAKTEATTAASLPKTSYGHYDDDDIPFPGINDDDPKVQDDDDTDSQ
DTTTPDVVDDPGSGYQSYSGNNAPDLVLFDLDEDEDKTPVNRSTKGGQQ
KNSQGHIEGRQTSRPIONVPGPHRTIHASAPLTDNRRNEPSGSTSPMLTPIIN
EADPLDADDETSLLPLESDDDEQDSDGTNRTPVAPPAVYRDHSEKKELPQDE
QQDQHYTQBEARNQSDNTQSEHSEFEEMRYHILRSQGPFDVAVLYHMKDEPVVFTSD
GKQYTPDLSLEEEYPPMLTEKAMNEENRFVTLGQQQFYPVMNHKNKFMALLQHHQ"
/gene="NP"
3027..3031
/notes="intergenic region"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
/product="VP35"
3032..3043
/gene="VP35"
/notes="transcription start signal"
3129..4151
/gene="VP35"
/codon_start=1
/product="VP35"
/protein_id="AAW76032.1"
/db_xref="GI:21702649"
/translation="MTTRTKRGHTVATTNDRMPGPELSGWISQLMTGRIPVSDIF
CDIENNPGLCYASQMQTKPNPKTRNSQTQDPCINHSFEEVVTQSLATVVOQQTII
ASELSIORITSLNGLKPVYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAATE
AYWAHQPPGPRSLYEESAIRKTESRDETVPQSVREAFNNLSTTSLTEENFGKPD
ISAKDLRNIMDHLPGFTGAHQVLQVQIKGKSNLSLDIIHAERQASLAEGSDSQCA
LIQITKRVPIFQDAAPPVHIHRSRGDIPRACQKSLRPVPPSPVDRGWVCVFLQDQK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
/gene="VP40"
/product="membrane-associated matrix protein VP40"
4390..4401
/gene="VP40"
/notes="transcription start signal"
4397..4407
/gene="VP35"
4479..5459
/gene="VP40"
/codon_start=1
/product="membrane-associated matrix protein VP40"
/protein_id="AAW76033.1"
/db_xref="GI:21702650"
/translation="MRRVILPTAPPEYMEAIYPVRSNSTIARGGNSNTGFLTPEVSNG
DTPGNLPIADDTIDHASHTPGSVSSAFILEAMVNVISGPKVLMKQIPILWPLGVAD
QKTSYSDTSTAAILASYTIHFQKATNPLVRVNRGLFGCIPDHPFLRLRIGNQALQE
FVLVPQVLPQVFTFDLTALKITQPLPAATWTDTPGSGALRPGISFHPKRLTLL
PNKSGKGSADLTSPEKIQAIMTSLODPKIVPIDPTKNIMGIEVPEVLVHKLTKGV
TSKSGQPIPLVLLPKYIGLDVPAFGDLTWITVTDQCDTCHSPASLPAVIEK"
5883..5894
/gene="VP40"
5895..5899
/notes="intergenic region"
5900..8305
/gene="GP"
5900..8305
/gene="GP"
misc_signal 8483..8489
/misc_feature 8483..8489
gene 8483..8489
mRNA 8483..8489
```



```
Qy 781 TATACAAGTGGGAAAAAGGAGCAATACCGGGAATACTAATTTGGAAGGTCAACCCCGAA 840
Db 6819 TATACAAGTGGGAAAAAGGAGCAATACCGGGAATACTAATTTGGAAGGTCAACCCCGAA 6878
Qy 841 ATTGATACAACAAATCGGGAGTGGGCTTCTCGGAAACTAAAAAAACCTCACTAGAAA 900
Db 6879 ATTGATACAACAAATCGGGAGTGGGCTTCTCGGAAACT-AAAAAAACCTCACTAGAAA 6937
Qy 901 ATTTCGAGTGAAGAGTGTGCTTTTCAGTGTGTATCAACGGAGCCAAAACATCATGTGT 960
Db 6938 ATTTCGAGTGAAGAGTGTGCTTTTCAGTGTGTATCAACGGAGCCAAAACATCATGTGT 6997
Qy 961 CAGAGTCCGGCGCGAACTTTCTTCGACCGAGGACCAACAACAACCTGAAGACCAAAA 1020
Db 6998 CAGAGTCCGGCGCGAACTTTCTTCGACCGAGGACCAACAACAACCTGAAGACCAAAA 7057
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db 7058 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 7117
Qy 1081 GCAGTGTGCATCTAACAACCCCTTGCACAATCTCCACAGTCCCAATCCCTCAACAAC 1140
Db 7118 GCAGTGTGCATCTAACAACCCCTTGCACAATCTCCACAGTCCCAATCCCTCAACAAC 7177
Qy 1141 AAACCAAGTCCCGACAAACAGCACCCCAATAATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 7178 AAACCAAGTCCCGACAAACAGCACCCCAATAATACACCGGTGTATAAATTGACATCTCTGAG 7237
Qy 1201 GCACCTCAAGTTGACAAACATCCCGCAGAACAGACAGCAGCAGCAGCCTCCGACACT 1260
Db 7238 GCACCTCAAGTTGACAAACATCCCGCAGAACAGACAGCAGCAGCAGCCTCCGACACT 7297
Qy 1261 CCTCTGCGACAGCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 1320
Db 7298 CCTCTGCGACAGCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 7357
Qy 1321 ACTGACTTCTCTGACCCCGCCACCAACAAGTCCCAAAACACACAGCGAGACCGCTGCG 1380
Db 7358 ACTGACTTCTCTGACCCCGCCACCAACAAGTCCCAAAACACACAGCGAGACCGCTGCG 7417
Qy 1381 AACAAACACACTCATCACAAGATACCGGAGAGAGAGTCCGACGCGGGAGCTAGGC 1440
Db 7418 AACAAACACACTCATCACAAGATACCGGAGAGAGAGTCCGACGCGGGAGCTAGGC 7477
Qy 1441 TTAATTACCAATACTATTCTGAGTTCGAGCTGACGACTATCAGCGGGGAGAGAACTCGA 1500
Db 7478 TTAATTACCAATACTATTCTGAGTTCGAGCTGACGAGTATCAGCGGGGAGAGAACTCGA 7537
Qy 1501 CGA 1503
Db 7538 AGA 7540

RESULT 13
AY354458
LOCUS
DEFINITION Zaire ebolavirus strain Zaire 1995, complete genome.
ACCESSION AY354458
VERSION AY354458.1 GI:33860540
KEYWORDS
SOURCE
ORGANISM
Zaire ebolavirus (ZEBOV)
Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
1 (bases 1 to 18961)
Chain, P.S.G., Ichou, M.A., Malfatti, S.A., Hajjaj, A., Vergez, L.M.,
Paragas, J., Do, L.H., Jahrling, P.B., Smith, K.L., McCready, P.M. and
Ibrahim, M.S.
Zaire Ebola virus, complete genome
Unpublished
2 (bases 1 to 18961)
Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H.,
Smith, K.L. and McCready, P.M.
```

```
Direct Submission
Submitted (24-JUL-2003) Viral Sequencing Group, Lawrence Livermore
National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
3 (bases 1 to 18961)
REFERENCE
Ichou, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Lofets, L.,
Hevey, M. and Schmaljohn, A.
Direct Submission
Submitted (24-JUL-2003) Viral Genomics Group, USAMRIID, Fort
Detrick, MD 21702, USA
FEATURES
Location/Qualifiers
1..18961
/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Zaire 1995"
/db_xref="taxon:186538"
1..55
/note="putative leader region"
/function="regulation or initiation of RNA replication"
56..3026
/gene="NP"
56..3026
/gene="NP"
/product="nucleoprotein"
56..67
/misc_signal
/gene="NP"
/note="putative; transcription start signal"
470..2689
/gene="NP"
/function="encapsidation of genomic RNA"
/codon_start=1
/product="nucleoprotein"
/protein_id="AAQ55045.1"
/db_xref="GI:33860541"
/translation="MDSRPQKVMVTPSITSDMDVHKILTAGLSVQQGIVRVRIPVY
QVNLERICQLIIQAFENGVDQFSADSFLLMLCHHAYQDYKLFLESGAVKYLEGH
GFREVKKRDKVGRLELLPAVSSGKIKRTLAAMEEETEAENAGQLSFASFLPDK
LVGEKACLEKVRQIQVHAEQGLIQVTAQSQVGHMMVIFRLMRTNFIKPLLIHQG
NMTVAGHDANDAVISNSVAQARFSGLLIVTKLVHLQKTERGVRLHPLARTAKVNE
VNSFAKLSLAKHGEVAPFARLLNSGVNLEHGLFQLSAIALGVATAGSTLAGV
NVGQYQOLREAATEAEKQLOQVAESRELHGLDDQEKILMNFQKNETSFQNTN
AMVILKERLAKLTEATASLSPTKTSYHYDDDDIIPFPINDDNPNQDQDDTDSQ
DTTIPDVVDDPDGSGEYSYSENGMAPDDLVLFDJEDDEDKTPVNRSTKGGQK
KNSQKQKTEGRTQSRPTQNVPGPHRTIHASAPLTNDNRNEPSGSTSRMLTPIN
EADPLDDADDETSPLPESDBEDRGTSNRTPTVAPPAVYRDHSKRELPQDE
QODDHTQAEARNQSDNMTQPSHSFEEMVYRHLRSQGFDAVLYHHMKDEPVPVFTSD
GKEYTPDLSLEEYPPMLTEKANEENRFTLDGQQFYWPVMNHKMFMAILQHHQ"
3015..3026
/gene="NP"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
/product="VP35"
3032..3043
/gene="VP35"
/note="putative; transcription start signal"
3129..4151
/gene="VP35"
/function="polymerase complex protein"
/codon_start=1
/product="VP35"
/db_xref="GI:33860542"
/translation="MTTRTKRGHTAATTONDRMPGPELSGWISEQLMTGRTPVSDIP
CDIENPGLCYASQMKQPKPKPKTNSQTQDPCINHSFEEVQVTLASLATVQQQTI
ASESLQKITSLENGLEKPKPKPKTNSQTQDPCINHSFEEVQVTLASLATVQQQTI
AYSAHGGQPPGPPLESIAESAIRKIESRDETVQSVREAFNNLDSTSLTEENFKQPO
ISAKDLRNMIDYHLPGFCTAFHQLVQVTKLGKDSNLDIIHAEFQASLAEDSPQCA
LIQITKRVPIFDAAAPPVHIRSRGDIIPRACQKSLRPVPPSPKIDRGWVCPVQLQDGK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
```

```

/misc_signal
/genes="VP40"
/product="matrix protein"
4390..4401
/genes="VP35"
/notes="transcription start signal"
4397..4407
/genes="VP35"
4479..5459
/genes="VP40"
/codon_start=1
/product="matrix protein"
/protein_id="AAQ55047.1"
/db_xref="GI:33860543"
/translation="MRRVILTAPPEYMEALYFVRSNSTIARGNSNTGFLTPPSVNG
DTPSPRLPIADDTIDHASHTPGSVSSAFLEAMVNVISGPKVLMMQIPIWLPLGVAD
FVLPFDLPQVFTDLTALKLTITPLPAATWTDPTGSGALRPGISFHPKLRPILL
PNKSGKNSADLTSPEKIQAIMTSLODFKIVPIDPTKNIINGIEVPETLVHKLTKKV
TSKNGQPIIPVLLPKYIGLDVPAQGLTMVITQDCDCHSPASLPVIEK"
5883..5894
/genes="VP40"
/genes="VP40"
5900..8306
/genes="GP"
5900..8306
/genes="GP"
/product="virion spike glycoprotein precursor"
/notes="uneditd mRNA"
join(5900..6923,6925..8306)
/genes="GP"
/product="small non-structural secreted glycoprotein"
/notes="RNA editing"
join(5900..6923,6926..8306)
/genes="GP"
/product="second non-structural secreted glycoprotein"
/notes="RNA editing"
5900..5911
/genes="GP"
/notes="putative; transcription start signal"
6039..8069
/genes="GP"
/function="receptor binding and fusion"
/notes="encodes two disulfide linked subunits GP1 and GP2"
/codon_start=1
/product="virion spike glycoprotein precursor"
/protein_id="AAQ55048.1"
/db_xref="GI:33860544"
/translation="MGVTGILQLPRDRFKRTSFFLWVILFORTFISIPLGVIHNSLTQ
VSEVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKRWGFRSGVPPKVVNYEAG
EWAENCYNLEIKKPDGSECLPAAPDGIRGPRCRYVHKVSGTGPCAGDFAFHKGAF
LYDRLASTVIYRGTTFAEYVAFILIPQAKKDFSSHPLEPNVATDPSSGYSTTI
RYQATGFTNETEYLFVDNLTYVQLESRFTPQLQLNETIYTSGRKSNITGKLIWK
VNPEDITTTIGEMAFWETKTKIRSELSTFVSNRAKNISGSPARTSSDPQNT
TTEDHKIWASNSAMVOVHSGREAAVSHLTATISTSPQPTTKPGDPSNTHPT
VYKLDISATQVEQHRITDNDSTASDTPPATTAAGPPEKAENTWSKTDLLDPATTT
SPQNSHETANNTHHQDTGESASSGKGLITNTIAGVAGLITGRRRAREAIINAQ
PKCNPLHYWTTQDEGAAIGLAWIPYFGFAEYITGLMHNQDGLICGLRLANETT
QALQLFLRATRTFTSLNRKADFLQLRWGGTCHILGPDCCIEPHDWTNITDKID
QIHFVDVKTLPDQGDNDNWTGWRQIPAGIVGTGVIIVIALFICICKFV"
7530..7541
/genes="GP"
/notes="cleaved by subtilisin-like cellular protease furin
into subunits GP1 and GP2 that are linked by a disulfide
bond; cleavage site"
7794..7871
/genes="GP"
/notes="Region: immunosuppressive motif"
7989..8054
/genes="GP"
/notes="transmembrane anchor; transmembrane-region site"
join(6039..6923,6925..7134)
/genes="GP"
/notes="sGP; secreted as a anti-parallel oriented
homodimer"

```

```

/codon_start=1
/exception="RNA editing"
/product="small non-structural, secreted glycoprotein"
/protein_id="AAQ55049.1"
/db_xref="GI:33860545"
/translation="MGVTGILQLPRDRFKRTSFFLWVILFORTFISIPLGVIHNSLTQ
VSEVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKRWGFRSGVPPKVVNYEAG
EWAENCYNLEIKKPDGSECLPAAPDGIRGPRCRYVHKVSGTGPCAGDFAFHKGAF
LYDRLASTVIYRGTTFAEYVAFILIPQAKKDFSSHPLEPNVATDPSSGYSTTI
RYQATGFTNETEYLFVDNLTYVQLESRFTPQLQLNETIYTSGRKSNITGKLIWK
VNPEDITTTIGEMAFWETKTKIRSELSTFVSNRAKNISGSPARTSSDPQNT
TTEDHKIWASNSAMVOVHSGREAAVSHLTATISTSPQPTTKPGDPSNTHPT
VYKLDISATQVEQHRITDNDSTASDTPPATTAAGPPEKAENTWSKTDLLDPATTT
SPQNSHETANNTHHQDTGESASSGKGLITNTIAGVAGLITGRRRAREAIINAQ
PKCNPLHYWTTQDEGAAIGLAWIPYFGFAEYITGLMHNQDGLICGLRLANETT
QALQLFLRATRTFTSLNRKADFLQLRWGGTCHILGPDCCIEPHDWTNITDKID
QIHFVDVKTLPDQGDNDNWTGWRQIPAGIVGTGVIIVIALFICICKFV"
5883..5894
/genes="VP40"
/genes="VP40"
5900..8306
/genes="GP"
5900..8306
/genes="GP"
/product="virion spike glycoprotein precursor"
/notes="uneditd mRNA"
join(5900..6923,6925..8306)
/genes="GP"
/product="small non-structural secreted glycoprotein"
/notes="RNA editing"
join(5900..6923,6926..8306)
/genes="GP"
/product="second non-structural secreted glycoprotein"
/notes="RNA editing"
5900..5911
/genes="GP"
/notes="putative; transcription start signal"
6039..8069
/genes="GP"
/function="receptor binding and fusion"
/notes="encodes two disulfide linked subunits GP1 and GP2"
/codon_start=1
/product="virion spike glycoprotein precursor"
/protein_id="AAQ55048.1"
/db_xref="GI:33860544"
/translation="MGVTGILQLPRDRFKRTSFFLWVILFORTFISIPLGVIHNSLTQ
VSEVDKLVCRDKLSSTNQLRSVGLNLEGNVATDVPSATKRWGFRSGVPPKVVNYEAG
EWAENCYNLEIKKPDGSECLPAAPDGIRGPRCRYVHKVSGTGPCAGDFAFHKGAF
LYDRLASTVIYRGTTFAEYVAFILIPQAKKDFSSHPLEPNVATDPSSGYSTTI
RYQATGFTNETEYLFVDNLTYVQLESRFTPQLQLNETIYTSGRKSNITGKLIWK
VNPEDITTTIGEMAFWETKTKIRSELSTFVSNRAKNISGSPARTSSDPQNT
TTEDHKIWASNSAMVOVHSGREAAVSHLTATISTSPQPTTKPGDPSNTHPT
VYKLDISATQVEQHRITDNDSTASDTPPATTAAGPPEKAENTWSKTDLLDPATTT
SPQNSHETANNTHHQDTGESASSGKGLITNTIAGVAGLITGRRRAREAIINAQ
PKCNPLHYWTTQDEGAAIGLAWIPYFGFAEYITGLMHNQDGLICGLRLANETT
QALQLFLRATRTFTSLNRKADFLQLRWGGTCHILGPDCCIEPHDWTNITDKID
QIHFVDVKTLPDQGDNDNWTGWRQIPAGIVGTGVIIVIALFICICKFV"
7530..7541
/genes="GP"
/notes="cleaved by subtilisin-like cellular protease furin
into subunits GP1 and GP2 that are linked by a disulfide
bond; cleavage site"
7794..7871
/genes="GP"
/notes="Region: immunosuppressive motif"
7989..8054
/genes="GP"
/notes="transmembrane anchor; transmembrane-region site"
join(6039..6923,6925..7134)
/genes="GP"
/notes="sGP; secreted as a anti-parallel oriented
homodimer"

```

```

Query Match      65.0%; Score 1464.6; DB 14; Length 18961;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1479; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGGCGGTACAGGAATATTCAGTTACCTCGTATCGATTCAGTTCAGAGGACATCATCTTT 60
Db 6039 ATGGGTGTACAGGAATATTCAGTTACCTCGTATCGATTCAGTTCAGAGGACATCATCTTT 6098
Qy 61 CTTTGGGTAATATTCCTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCAATCCACAAT 120
Db 6099 CTTTGGGTAATATTCCTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCAATCCACAAT 6158
Qy 121 AGCACATTACAGGTAGTGTGCGACAACTAGTTTTCGTGACAACTGTGTCATCCACA 180
Db 6159 AGCACATTACAGGTAGTGTGCGACAACTAGTTTTCGTGACAACTGTGTCATCCACA 6218
Qy 181 AATCAATTGAGATCAGTTGGATGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 240
Db 6219 AATCAATTGAGATCAGTTGGATGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 6278
Qy 241 TCTCAACTAAAGATGGGGCTTCAGGTCCGGTCCCAACCAAGGTGGTCAATATATCAA 300
Db 6279 TCTCAACTAAAGATGGGGCTTCAGGTCCGGTCCCAACCAAGGTGGTCAATATATCAA 6338
Qy 301 GCTGTTGAATGGGTGAAACTGCTACAAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db 6339 GCTGTTGAATGGGTGAAACTGCTACAAATCTTGAATCAAAAAACCTGACGGGAGTGAG 6398
Qy 361 TGTCTACAGAGCCGACAGCGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 420
Db 6399 TGTCTACAGAGCCGACAGCGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 6458
Qy 421 GTATCAGAGACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGGTCTTCTTC 480
Db 6459 GTATCAGAGACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGGTCTTCTTC 6518

```

```
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db |||
QY 6519 CTGTATGACCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 6578
Db |||
QY 541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTTCAGCTCAGCCCCCTTGAGA 600
Db |||
QY 6579 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTTCAGCTCAGCCCCCTTGAGA 6638
QY 601 GAGCCGGTCAATGCAACGAGGAGACCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
Db |||
QY 6639 GAGCCGGTCAATGCAACGAGGAGACCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 6698
QY 661 CAGGTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db |||
QY 6699 CAAGCTACCGGTTTGGAAACCAATGAGACAGAGTATTGTTCGAGGTTGACAATTTGACC 6758
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
Db |||
QY 6759 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGTGAATGAGACAATA 6818
QY 781 TATACAAGTGGGAAAGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 840
Db |||
QY 6819 TATACAAGTGGGAAAGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 6878
QY 841 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 900
Db |||
QY 6879 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 6938
QY 901 ATTCCGAGTGAAGATTGTCTTTCACAGTTGTATCAACGGAGGCCAAAACATCATGTGT 960
Db |||
QY 6939 ATTCCGAGTGAAGATTGTCTTTCACAGTTGTATCAACAGAGGCCAAAACATCATGTGT 6998
QY 961 CAGAGTCCGGCGCGAACTTCTTCGACCCAGGAGCCACACACAACTGAAGACACACAA 1020
Db |||
QY 6999 CAGAGTCCGGCGCGAACTTCTTCGACCCAGGAGCCACACAACTGAAGACACACAA 7058
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db |||
QY 7059 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 7118
QY 1081 GCAAGTGTGCGATCTAACAACCTTGGCCACAATCTCCACAGTCCCAATTCCTCTCAACCC 1140
Db |||
QY 7119 GCAAGTGTGCGATCTAACAACCTTGGCCACAATCTCCACAGTCCCAATTCCTCTCAACCC 7178
QY 1141 AAACAGGTCCGGACAAACAGCCCAATATACACCGCTGTATAAATCTTGACATCTCTGAG 1200
Db |||
QY 7179 AAACAGGTCCGGACAAACAGCCCAATATACACCGCTGTATAAATCTTGACATCTCTGAG 7238
QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAACAGAACAGACAGACAGCCCTCCGACACT 1260
Db |||
QY 7239 GCAACTCAAGTTGAACAACATCACCGCAGAACAGAACAGACAGACAGCCCTCCGACACT 7298
QY 1261 CCGTCTGCCACAGCCGACCGGACCCCAAGAGAGAGAACACCAACAGACAGGAGAGGAC 1320
Db |||
QY 7299 CCGCCGCGCACAGCCGACCGGACCCCAAGAGAGAGAACACCAACAGACAGGAGAGGCT 7358
QY 1321 ACTGACTTCTCGACCCCGCCACCAACAAGTCCCAAAACACAGCGAGACCCGCTGCG 1380
Db |||
QY 7359 ACCGACTTCTCGACCCCGCCACCAACAAGTCCCAAAACACAGCGAGACCCGCTGCG 7418
QY 1381 AACAAACACACTCATCACCAAGATACCGGAGAGAGAGTCCGACGAGCGGGAAGCTTAGGC 1440
Db |||
QY 7419 AACAAACACACTCATCACCAAGATACCGGAGAGAGAGTCCGACGAGCGGGAAGCTTAGGC 7478
QY 1441 TTAATTACCAATATCTTCTGAGTCCGAGGACTGATCACAGGCGGAGAGAACTCGA 1500
Db |||
QY 7479 TTAATTACCAATATCTTCTGAGTCCGAGGACTGATCACAGGCGGAGAGAGCTCGA 7538
QY 1501 CGA 1503
Db 7539 AGA 7541
```

```
RESULT 14
EVU28077
LOCUS
DEFINITION
Zaire Ebola virus strain Zaire95 virion spike glycoprotein (SP)
gene, complete cds, and small/secreted glycoprotein precursor (SGP)
gene, complete cds.
ACCESSION
U28077
VERSION
U28077.1
KEYWORDS
GI:1695252
SOURCE
Zaire ebolavirus (ZEBOV)
ORGANISM
Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE
1 (bases 1 to 2408)
Sanchez,A., Trappier,S.G., Mahy,B.W., Peters,C.J. and Nichol,S.T.
The virion glycoproteins of Ebola viruses are encoded in two
reading frames and are expressed through transcriptional editing
Proc. Natl. Acad. Sci. U.S.A. 93 (8), 3602-3607 (1996)
96195018
PUBMED
8622982
REFERENCE
2 (bases 1 to 2408)
Sanchez,A.
Direct Submission
Submitted (31-MAY-1995) Anthony Sanchez, Special Pathogens Branch,
Division of Viral and Rickettsial Diseases, Centers for Disease
Control and Prevention, 1600 Clifton Road, Bldg. 15, Room SB611,
Atlanta, GA 30333, USA
COMMENT
On Nov 30, 1996 this sequence version replaced gi:1041220.
FEATURES
Location/Qualifiers
source
1..2408
/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Zaire95"
/db_xref="taxon:186538"
1..141
5'UTR
misc_signal
1..14
CDS
/translation="MGVTGILQLPRDRKRTSFFLWVILFORTFSPILGVHNSTLQ
VSDVKLCVRDKLSTNQLRSVGLNLEGVATDVPSTKRWGFRSGVPPKVVNVEAG
EWAECYNLEIKKPDGSECLPAAPDGIKRGPRCRYVHKVSGTGPCAGFAFHKEGAF
LYDLRASTVIYRGTTFAEGVVAFLILPOAKKDDFFSSHPLREFVNATEDPSSGYITII
RYQATGFTNTEYLFEVDNLTYYQLESRFTPQFLQLNETIYTSGKRSNTTGKLIWK
VNPEDITTTIGWAFWETKNTKRIRSELSFTAVSNRAKNIISGSPPTKSDPGTNT
TTEDHKIMASENSAMVQVHSGREAAVSHLTTLATISTSPQPTTKGPDNSTHTPT
VYKLDISPAQVEQHRRTDNDSTADTPPATTAAGPLKAENTNTSGKTDLLDPATTT
SPQNSHETAGNNTHQDTGESASGKGLITNTIAGVAGLITGGRARAEIYNAQ
PKCPNLRHYTTQDEGAAGIAPYFGPAABEGYIEGIMNQDGLICGLRQIANETT
QALQFLRATLFTSLNKAIDFLQRWGTCHILGPDCCICPHDWTNKITDKID
QIITDFVDKTLPLDQDNDNMTGWRQIPAGITGVITGVIAVIAFCICKFV"
142..1236
/gene="SGP"
142..1236
/gene="SGP"
142..1236
/function="unknown"
/note="nonstructural protein; editing event during
transcription occurs that links the 0 and -1 frames and
leads to the formation of the spike (surface) structure on
the virion"
/codon_start=1
/product="small/secreted glycoprotein precursor"
/protein_id="AAB37094.1"
/db_xref="GI:1695252"
/translation="MGVTGILQLPRDRKRTSFFLWVILFORTFSPILGVHNSTLQ
VSDVKLCVRDKLSTNQLRSVGLNLEGVATDVPSTKRWGFRSGVPPKVVNVEAG
EWAECYNLEIKKPDGSECLPAAPDGIKRGPRCRYVHKVSGTGPCAGFAFHKEGAF
LYDLRASTVIYRGTTFAEGVVAFLILPOAKKDDFFSSHPLREFVNATEDPSSGYITII
```

```
RYATGTCGTNETEYLFEVDNLTYVOLESFPTPOFLLOLNETIYTSKESNTTGKLIWK
VNPELDTTIGWAFWETKTSLEKFAVKSLSQLYQTEPKTSVVRRELLPTQGPQTQ
QUKTKSWIKIPLQWFKCTVKEGKIQCRI"
142..237
/gene="SGP"
238..1230
/gene="SGP"
/product="small/secreted glycoprotein"
1021..1027
/gene="SGP"
/note="transcription editing site"
2088..2156
/gene="GP"
2088..2156
/gene="GP"
/note="encodes transmembrane anchor region"
2172..2408
2398..2408
ORIGIN

Query Match      64.5%; Score 1452.6; DB 14; Length 2408;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTCTTT 60
Db 142 ATGGGTGTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCTCTTT 201

QY 61 CTTTGGGTAATATTCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 202 CTTTGGGTAATATTCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 261

QY 121 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 180
Db 262 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 321

QY 181 AATCAATTTGAGATCAGTTGGATCTCGAAGGGAATGAGTGGCACTGACGTGCA 240
Db 322 AATCAATTTGAGATCAGTTGGATCTCGAAGGGAATGAGTGGCACTGACGTGCA 381

QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCACCAAGGTGGTCAATTTATCAA 300
Db 382 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCACCAAGGTGGTCAATTTATGAA 441

QY 301 GCTGTGTAATGGGCTGAAAACCTGCTACAATCTTGAATCAAAAACTGACGGGAGTGAG 360
Db 442 GCTGTGTAATGGGCTGAAAACCTGCTACAATCTTGAATCAAAAACTGACGGGAGTGAG 501

QY 361 TGTCTACGAGCAGCGCCAGACGGGATTCGGGCTTCCCGCGTGCCTGATGTGCACAAA 420
Db 502 TGTCTACGAGCAGCGCCAGACGGGATTCGGGCTTCCCGCGTGCCTGATGTGCACAAA 561

QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 480
Db 562 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 621

QY 481 CTGTATGATTCGACTTGTCTTCCAAGTTATCTACCGAGGAACGACTTTTCCTGAAGGTGTC 540
Db 622 CTGTATGATTCGACTTGTCTTCCAAGTTATCTACCGAGGAACGACTTTTCCTGAAGGTGTC 681

QY 541 GTTGATTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 682 GTTGATTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 741

QY 601 GAGCGGCTCAATGCAACGAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 742 GAGCGGCTCAATGCAACGAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801

QY 661 CAGGCTACCGGTTTGGAACTAATGAGACAGAGTACTTGTTCGAGGTTTGACAATTTGACC 720
Db 802 CAACTACCGGTTTGGAACTAATGAGACAGAGTACTTGTTCGAGGTTTGACAATTTGACC 861

QY 721 TACGTCCAACTTGAATCAAGATTTCACCAACAGTTTTCGTCTCCAGCTGAATGAGACAATA 780
```

```
Db 862 TACGTCCAACTTGAATCAAGATTTCACCAACAGTTTTCGTCTCCAGCTGAATGAGACAATA 921
QY 781 TATACAAGTGGGAAAAGAGCAATATACACGGGAAAACTAATTTTGGAAAGGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAAGAGCAATATACACGGGAAAACTAATTTTGGAAAGGTCAACCCCGAA 981
QY 841 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAAACTAAAAAACCTCACTAGAAAA 900
Db 982 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAACT-AAAAAACCTCACTAGAAAA 1040
QY 901 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAACGAGGACCAAAAAATCATAGTGT 960
Db 1041 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAACGAGGACCAAAAAATCATAGTGT 1100
QY 961 CAGAGTCCGGCGGAACTTTCTTCGAGCCCGAGGACCAACAACTGAAGAGACCAAAA 1020
Db 1101 CAGAGTCCGGCGGAACTTTCTTCGAGCCCGAGGACCAACAACTGAAGAGACCAAAA 1160
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAAAGCT 1080
Db 1161 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAAAGCT 1220
QY 1081 GCAGTGTGCACTTAAACAACCTTTGCCACAACTCTCCAGAGTCCCAATCCCTCAACAACC 1140
Db 1221 GCAGTGTGCACTTAAACAACCTTTGCCACAACTCTCCAGAGTCCCTCAACAACC 1280
QY 1141 AAACAGGTCCGGCAACAGCACCCCAATAATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 1281 AAACAGGTCCGGCAACAGCACCCCAATAATACACCGGTGTATAAATTGACATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAACAACTACCGCGAGAAACAGAAACGACAGACAGCTTCGACACT 1260
Db 1341 GCAACTCAAGTTGAACAACTACCGCGAGAAACAGAAACGACAGACAGCTTCGACACT 1400
QY 1261 CCCTCTGCACGACCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC 1320
Db 1401 CCCCCGCGACGACCGGACCGGACCCCTAAAGAGAGAAACCAACACGAGCAAGGGT 1460
QY 1321 ACTGACTTCTGAGCCCGCCACCAACAACTAGTCCCAAAACCAACAGGAGACCGCTGCGC 1380
Db 1461 ACCGACCTCTGAGCCCGCCACCAACAACTAGTCCCAAAACCAACAGGAGACCGCTGCGC 1520
QY 1381 AACAAACACTCATCAACAAAGATAACGAGAGAGAGTGCAGAGCGGAGAGTACGAGC 1440
Db 1521 AACAAACACTCATCAACAAAGATAACGAGAGAGAGTGCAGAGCGGAGAGTACGAGC 1580
QY 1441 TTAATTACCACTATTTGCTGAGTGCAGGACTGATCAAGGCGGAGAGAGTACGAGC 1500
Db 1581 TTAATTACCACTATTTGCTGAGTGCAGGACTGATCAAGGCGGAGAGAGTACGAGC 1640
QY 1501 CGA 1503
Db 1641 AGA 1643

RESULT 15
EVU77384
LOCUS EVU77384
DEFINITION Zaire Ebola virus strain Gabon-94 virion spike glycoprotein (GP)
ACCESSION U77384
VERSION U77384.1
KEYWORDS GI:2138276
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (bases 1 to 2446)
AUTHORS Volchkov,V.E., Blinov,V.M., Kotov,A., Chepurnov,A. and Netesov,S.
TITLE The full-length nucleotide sequence of the Ebola virus
JOURNAL Thesis (1993) Institut f. Virologie, Philipps-Universitaet
REFERENCE 2 (bases 1 to 2446)
```

EVU77384 2446 bp RNA linear VRL 28-AUG-2002
Zaire Ebola virus strain Gabon-94 virion spike glycoprotein (GP)
U77384
U77384.1 GI:2138276

Qy	1021	ATCATGGCTTCAGAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAGGAGCT	1080
Db	1161	ATCATGGCTTCAGAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAGGAGCT	1220
Qy	1081	GCAGTGTGCGCATCTAACAAACCTTGCCCAATCTCCACGAGTCCCAATCCCTCACAACC	1140
Db	1221	GCAGTGTGCGCATCTGACAACCTTGCCCAATCTCCACGAGTCTTCGACCCCCCATACC	1280
Qy	1141	AAACGAGTCCGGACAACAGCAACCCATTAATACACCGTGTATATACTTGACATCTCTGAG	1200
Db	1281	AAACGAGTCCGGACAACAGCAACCCATTAATACACCGTGTATATACTTGACATCTCTGAG	1340
Qy	1201	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACGACGACACAGCCTCCGACACT	1260
Db	1341	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACGACGACACACCTCCGACACT	1400
Qy	1261	CCCTCTGCCACGACCGCGACCGGACCCCAAGCAGAGAACCAACACGAGCAAGAGC	1320
Db	1401	CCCTCTGCCACGACCGCGACCGGACCCCAAGCAGAGAACCAACACGAGCAAGAGC	1460
Qy	1321	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAAGCGAGACCGCTGGC	1380
Db	1461	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAAGCGAGACCGCTGGC	1520
Qy	1381	AACACACACTCATCACCAAGATACCGGAGAGAGTGCCAGCAGCGGAGGCTAGGC	1440
Db	1521	AACACACACTCATCACCAAGATACCGGAGAGAGTGCCAGCAGCGGAGGCTAGGC	1580
Qy	1441	TTAATTACCAATPACTATTGCTGGAGTCGAGGACTGATCACAGCGGAGAGAACTCGA	1500
Db	1581	TTAATTACCAATPACTATTGCTGGAGTCGAGGACTGATCACAGCGGAGAGAACTCGA	1640
Qy	1501	CGA 1503	
Db	1641	AGA 1643	

Search completed: August 1, 2005, 02:43:36
Job time : 9684.51 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1179.79 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-1

Perfect score: 2252

Sequence: 1 atgggcgtacaggaattt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N Geneseq 16Dec04:*
- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	2252	10	Abt13450 DNA encod
2	1506	66.9	2039	10	Abt13455 DNA encod
3	1501.4	66.7	2172	4	Aad04041 Ebola vir
4	1501.4	66.7	2224	3	Aaz51038 Ebola vir
5	1501.4	66.7	2298	3	Aaz87189 Ebola vir
6	1501.4	66.7	2298	4	Aaf76953 Ebola vir
7	1501.4	66.7	2298	12	Adl27460 Nucleotid
8	1501.4	66.7	6887	10	Acc71526 VRC6005 (
9	1501.4	66.7	6314	10	Acc71529 VRC6008 (
10	1501.4	66.7	7044	10	Acc71527 VRC6006 (
11	1501.4	66.7	7106	10	Acc71528 VRC6007 (
12	1501.4	66.7	7154	10	Acc71521 VRC6000 (
13	1501.4	66.7	7188	10	Acc71522 VRC6001 (
14	1501.4	66.7	8199	10	Acc71549 VRC6604 (
15	1501.4	66.7	8439	10	Acc71548 VRC6603 (
16	1501.4	66.7	10783	10	Acc71545 VRC6600 (
17	1499.8	66.6	7285	10	Aax59392 Plasmid p
18	1499.8	66.6	9106	12	Adf85419 Ebola vir
19	1495	66.4	7285	2	Aax89797 DNA of pv
20	1489.4	66.1	18959	13	Adm48337 Zaire ebo

21	1489.4	66.1	18959	13	Adm48364	Adm48364 Zaire ebo
22	1489.4	66.1	18959	13	Adm48355	Adm48355 Zaire ebo
23	1487.8	66.1	7272	2	Aax59393	Aax59393 Plasmid p
24	1484.6	65.9	7272	2	Aax89798	Aax89798 DNA of pv
25	1480	65.7	6724	10	Acc71525	Acc71525 VRC6004 (
26	906.4	40.2	6624	10	Acc71523	Acc71523 VRC6002 (
27	905.4	40.2	6561	10	Acc71524	Acc71524 VRC6003 (
28	900	40.0	6324	10	Acc71557	Acc71557 CMV/R Ebo
29	753	33.4	2051	10	Abt13454	Abt13454 DNA encod
30	746.6	33.2	2046	10	Abt13453	Abt13453 DNA encod
31	742.2	33.0	2100	3	Aaz51039	Aaz51039 Marburg v
32	742.2	33.0	2103	8	Acc57764	Acc57764 Marburg v
33	742.2	33.0	7778	10	Acc71550	Acc71550 VRC6701 (
34	742.2	33.0	11460	3	Aaz87211	Aaz87211 VEE repli
35	740.2	32.9	2164	4	Aad04042	Aad04042 Marburg v
36	685.8	30.5	19112	13	Adm48345	Adm48345 Marburg v
37	675.2	30.0	6467	10	Acc71530	Acc71530 VRC6052 (
38	644.2	28.6	7005	10	Acc71551	Acc71551 VRC6702 (
39	644.2	28.6	8256	10	Acc71552	Acc71552 VRC6710 (
40	548.2	24.3	2046	10	Abt13456	Abt13456 DNA encod
41	546.6	24.3	2046	10	Abt13452	Abt13452 DNA encod
42	546.2	24.3	6885	10	Acc71538	Acc71538 VRC6302 (
43	546.2	24.3	7002	10	Acc71536	Acc71536 VRC6300 (
44	546.2	24.3	7023	10	Acc71541	Acc71541 VRC6351 (
45	546.2	24.3	7036	10	Acc71537	Acc71537 VRC6301 (

ALIGNMENTS

RESULT 1
ABT13450
ID ABT13450 standard; DNA; 2252 BP.
XX
AC ABT13450;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein EBOV-GP1/MGV-GP2.
XX
KW Virutide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
DR P-PSDB; ABJ18472.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Claim 12; Page 66-68; 94pp; English.
XX
CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus

[illegible]

```
Db 2041 GACTTGTCAAAATAATTTTCAGAGCAAAATTCACCAATTTAAAAGGACGACAAAAGAG 2100
Qy 2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
Db 2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
Qy 2161 TTGGGCATTTTGGCTACTATTATTCATAGCTCTCTGTGATTGCTCTATCTCTGTATTGTGCT 2220
Db 2161 TTGGGCATTTTGGCTACTATTATTCATAGCTCTCTGTGATTGCTCTATCTCTGTATTGTGCT 2220
Qy 2221 ATCTTTACTAAATATATCGGATAACGGAAATTC 2252
Db 2221 ATCTTTACTAAATATATCGGATAACGGAAATTC 2252

RESULT 2
ABT13455
ID ABT13455 standard; DNA; 2039 BP.
XX
AC ABT13455;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein EBOV-GP1/GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
DR P-PSDB; ABJ18477.
XX
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Disclosure; Page 85-88; 94pp; English.
XX
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
XX
SQ Sequence 2039 BP; 613 A; 509 C; 466 G; 451 T; 0 U; 0 Other;

Query Match 66.9%; Score 1506; DB 10; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGTTACAGGAATATTGCAAGTACCTGTCGATCGAATTCAGAGGACATCATCTTTT 60
Db 1 ATGGGCGTTACAGGAATATTGCAAGTACCTGTCGATCGAATTCAGAGGACATCATCTTTT 60
Qy 61 CTTTGGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 61 CTTTGGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
```

```
Qy 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTCTCATCCACA 180
Db 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTCTCATCCACA 180
Qy 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Qy 241 TCTGCAACTAAAAGATGGGCTTTCCAGTCCGGTGTCCACCAAAAGGTGGTCAATTATGAA 300
Db 241 TCTGCAACTAAAAGATGGGCTTTCCAGTCCGGTGTCCACCAAAAGGTGGTCAATTATGAA 300
Qy 301 GCTGCTGAATGGGCTGAAAACTGCTACAATCTTGAATTCAAAAACCTGACGGAGTGAG 360
Db 301 GCTGCTGAATGGGCTGAAAACTGCTACAATCTTGAATTCAAAAACCTGACGGAGTGAG 360
Qy 361 TGTCTACACAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCCGGTATGTCACAAA 420
Db 361 TGTCTACACAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCCGGTATGTCACAAA 420
Qy 421 GTATCAGAAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 421 GTATCAGAAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Qy 481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTCGCTGGAAGGTGTC 540
Db 481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTCGCTGGAAGGTGTC 540
Qy 541 GTTGCAATTTCTGATACCTGCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 541 GTTGCAATTTCTGATACCTGCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 600
Qy 601 GAGCCGGTCAATGCAACGGAGGACCCGCTAGTGGCTACTATTCTACCACAATTAGATAT 660
Db 601 GAGCCGGTCAATGCAACGGAGGACCCGCTAGTGGCTACTATTCTACCACAATTAGATAT 660
Qy 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTGACAATTTGACC 720
Db 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTGACAATTTGACC 720
Qy 721 TACGTCCAACTTGAATCAAGATTCACACCACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 721 TACGTCCAACTTGAATCAAGATTCACACCACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Qy 781 TATACAAAGTGGGAAAAGGAGCAATACCAACGGGAAAACCTAATTTGGAGGTCAACCCGAA 840
Db 781 TATACAAAGTGGGAAAAGGAGCAATACCAACGGGAAAACCTAATTTGGAGGTCAACCCGAA 840
Qy 841 ATTGATACAAATCGGGGAGTGGGCTTCTGGGAAACTTAAAAAAACCTCACTAGAAAA 900
Db 841 ATTGATACAAATCGGGGAGTGGGCTTCTGGGAAACTTAAAAAAACCTCACTAGAAAA 900
Qy 901 ATTCGCAAGTGAAGAGTGTCTTTCACAGTTGCTATCAACCGGAGCCAAAACATCAGTGT 960
Db 901 ATTCGCAAGTGAAGAGTGTCTTTCACAGTTGCTATCAACCGGAGCCAAAACATCAGTGT 960
Qy 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACAACTGAAGAGACCAAAA 1020
Db 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACAACTGAAGAGACCAAAA 1020
Qy 1021 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db 1021 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Qy 1081 GCAGTGTGCAATCTAAACCCCTTGGCCACATCTCCAGGTGCTCCCTCATCCACACC 1140
Db 1081 GCAGTGTGCAATCTAAACCCCTTGGCCACATCTCCAGGTGCTCCCTCATCCACACC 1140
Qy 1141 AAACAGGTCCGGACAAACAGCACCCCAATAACACCCGCTGATAAACTTGACATCTCTGAG 1200
Db 1141 AAACAGGTCCGGACAAACAGCACCCCAATAACACCCGCTGATAAACTTGACATCTCTGAG 1200
```

```
QY 1201 GCAACTAAGTTGAACAACATCACCGCAGAACAGACAACAGACAGACAGCTCCGACACT 1260
Db 1201 GCAACTAAGTTGAACAACATCACCGCAGAACAGACAACAGACAGACAGCTCCGACACT 1260
QY 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAGAGAGAGAGACCAACACGAGGAGAGC 1320
Db 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAGAGAGAGAGACCAACACGAGGAGAGC 1320
QY 1321 ACTGACTTCTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAAGAGAGAGAGCCTGGC 1380
Db 1321 ACTGACTTCTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAAGAGAGAGCCTGGC 1380
QY 1381 AACAAACACATCATCACAAGATACCGGAGAGAGAGTCCAGACGCGGAGAGTATAGC 1440
Db 1381 AACAAACACATCATCACAAGATACCGGAGAGAGAGTCCAGACGCGGAGAGTATAGC 1440
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCGAGACTGATCAGAGCGGGAGAGACTCGA 1500
Db 1441 TTAATTACCAATACTATTGCTGGAGTCCGAGACTGATCAGAGCGGGAGAGACTCGA 1500
QY 1501 CGATCG 1506
Db 1501 CGATCG 1506

RESULT 3
AAD04041
ID AAD04041 standard; DNA; 2172 BP.
XX AAD04041;
AC AAD04041;
XX
DT 02-JUL-2001 (first entry)
XX
DE Ebola virus strain Zaire 95 glycoprotein (GP) DNA.
XX
KW Glycoprotein; GP; immune response; vaccine; antiviral; ds.
XX
OS Ebola virus.
XX
FH Key Location/Qualifiers
FT 142..2172
FT CDS /tag= a
FT /product= "Ebola virus strain Zaire 95 glycoprotein (GP)"
XX
XX US6200959-B1.
XX
XX 13-MAR-2001.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX (POWD-) POWDERJECT VACCINES INC.
XX
XX Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX
XX WPI: 2001-280564/29.
XX
XX P-PSDB; AAE00706.
XX
XX Inducing an immune response to a Marburg or Ebola virus involves
XX delivering a genetic vaccine, which contains a genetic construct encoding
XX antigenic determinants for filovirus, using a particle acceleration
XX device.
XX
XX Claim 4; Col 17-22; 33pp; English.
XX
XX The patent discloses a method of inducing an immune response to Marburg
XX or Ebola virus glycoprotein which involves delivering a genetic vaccine,
XX containing a genetic construct encoding antigenic determinants for
XX filovirus, using a particle acceleration device. The genetic vaccine for
XX filovirus is created by joining a DNA sequence encoding at least a
XX portion of the filovirus glycoprotein to a promoter effective to promote
XX transcription of the DNA sequence. This method is useful for inducing
```

```
CC humoral, cell-mediated and secretory immune responses in the treated
CC individual. The present sequence is a DNA encoding Ebola virus strain
CC Zaire 95 glycoprotein (GP), which is terminally siolated. This sequence
CC is used in the construction of genetic vaccine against filoviruses
XX
SQ Sequence 2172 BP; 656 A; 534 C; 496 G; 486 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 4; Length 2172;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATTGTCAGTTTACCTCGTGTGATTCGATTCAGAGGACATCATCTTT 60
Db 142 ATGGCGGTTACAGGAATATTGTCAGTTTACCTCGTGTGATTCGATTCAGAGGACATCATCTTT 201
QY 61 CTTTGGGTAATATTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACCAAT 120
Db 202 CTTTGGGTAATATTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACCAAT 261
QY 121 AGCACATTACAGGTTAGTGTGTCGACAAACTAGTTTTGTGTCGACAAACTGTGTCATCCACA 180
Db 262 AGCACATTACAGGTTAGTGTGTCGACAAACTAGTTTTGTGTCGACAAACTGTGTCATCCACA 321
QY 181 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 240
Db 322 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 381
QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTTCGCCACCAAGGTGGTCAATATATGAA 300
Db 382 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTTCGCCACCAAGGTGGTCAATATATGAA 441
QY 301 GCTGTGTAATGGGCTGAAAACCTGCTTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db 442 GCTGTGTAATGGGCTGAAAACCTGCTTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 501
QY 361 TGTCTACAGACGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCACAAA 420
Db 502 TGTCTACAGACGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCACAAA 561
QY 421 GTATCAGGAACGGGACCGTGTGCCGGAGACTTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 562 GTATCAGGAACGGGACCGTGTGCCGGAGACTTTTGCTTCCATAAAGAGGGTCTTTCTTC 621
QY 481 CTGTATGATCGACTTTGCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db 622 CTGTATGATCGACTTTGCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 681
QY 541 GTTGCAATTTCTGATCTACTGCCCAAGCTAAGAAAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 682 GTTGCAATTTCTGATCTACTGCCCAAGCTAAGAAAGGACTTCTTCAGCTCACACCCCTTGAGA 741
QY 601 GAGCCGGTCAATGCAACGGAGGACCGCTCTAGTGGCTACTATTCTACCACAATTTAGATAT 660
Db 742 GAGCCGGTCAATGCAACGGAGGACCGCTCTAGTGGCTACTATTCTACCACAATTTAGATAT 801
QY 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTTCCAGGTTTGACAATTTGACC 720
Db 802 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTTCCAGGTTTGACAATTTGACC 861
QY 721 TACGTTCCAACTTGAATCAAGATTTCACACACAGTTTTCGCTCCAGCTGATGAGACAAATA 780
Db 862 TACGTTCCAACTTGAATCAAGATTTCACACACAGTTTTCGCTCCAGCTGATGAGACAAATA 921
QY 781 TATACAAGTGGGAAAGGAGCAATACCACGGGAAACCTAATTTGGAAGGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAGGAGCAATACCACGGGAAACCTAATTTGGAAGGTCAACCCCGAA 981
QY 841 ATTGATACAAACAATCGGGAGTGGGCTTCTGGGAAACCTTAAAAAAACCTCTACAGAAA 900
Db 982 ATTGATACAAACAATCGGGAGTGGGCTTCTGGGAAACCTTAAAAAAACCTCTACAGAAA 1041
QY 901 ATTCCAGTGAAGATTGCTTTTTCAGTTTATCAAGTTGATCAACGGGAGCCAAAAACATCATGTGT 960
Db 1041 ATTCCAGTGAAGATTGCTTTTTCAGTTTATCAAGTTGATCAACGGGAGCCAAAAACATCATGTGT 960
```

```
Db 1042 ATTCCAGTGAAGATTGCTTTTCACAGTTGTATCAAAACGAGGCCAAAAACATCAGTGGT 1101
Qy 961 CAGAGTCGGCGCGAATCTTTCGGACCCGAGGCCAACACAACTGAAGACACAAA 1020
Db 1102 CAGAGTCGGCGCGAATCTTTCGGACCCGAGGCCAACACAACTGAAGACACAAA 1161
Qy 1021 ATCATGGCTTCAGAAATTCCTCTGCATGTTTCAAGTGCACAGTCAGGAGGAGCT 1080
Db 1162 ATCATGGCTTCAGAAATTCCTCTGCATGTTTCAAGTGCACAGTCAGGAGGAGCT 1221
Qy 1081 GCAGTGTGCGCATCTAACAACTTCGCCACAATCTCCAGAGTCCCAATCCCTCACACC 1140
Db 1222 GCAGTGTGCGCATCTAACAACTTCGCCACAATCTCCAGAGTCCCAATCCCTCACACC 1281
Qy 1141 AAACAGTTCGGACAAACAGCACCCATTAATACCCGTTGTATAAATTGACATCTCTGAG 1200
Db 1282 AAACAGTTCGGACAAACAGCACCCATTAATACCCGTTGTATAAATTGACATCTCTGAG 1341
Qy 1201 GCAACTCAAGTTGAACACATCACCGCAGAACAGACACAGCAGACAGCTCCGACACT 1260
Db 1342 GCAACTCAAGTTGAACACATCACCGCAGAACAGACACAGCAGACAGCTCCGACACT 1401
Qy 1261 CCTCTGCGACGACGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAAGAGC 1320
Db 1402 CCTCTGCGACGACGCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAAGAGC 1461
Qy 1321 ACTGACTTCTGGACCCCGCACCAACAAAGTCCCAAAACCCAGCAGACCGCTGGC 1380
Db 1462 ACTGACTTCTGGACCCCGCACCAACAAAGTCCCAAAACCCAGCAGACCGCTGGC 1521
Qy 1381 AACAAACACTCATCACAGATHACCGGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1440
Db 1522 AACAAACACTCATCACAGATHACCGGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1581
Qy 1441 TTAATTACCAATATCTGAGTCCGAGCTGATCAGAGCGGGAAGAACTCGA 1500
Db 1582 TTAATTACCAATATCTGAGTCCGAGCTGATCAGAGCGGGAAGAACTCGA 1641

RESULT 4
AAZ51038
ID AAZ51038 standard; DNA; 2224 BP.
XX
AC AAZ51038;
XX
DT 05-JUN-2000 (first entry)
XX
DE Ebola virus Zaire envelope glycoprotein DNA.
XX
KW Ebola-Zaire virus envelope glycoprotein; viral glycoprotein;
KW pseudotyped retrovirus; MMLV; Moloney murine leukaemia virus; antibody;
KW transduction; screening agent; immunological agent;
KW pharmacological agent; ss.
OS Ebola virus.
XX
FH Key Location/Qualifiers
FT CDS 11..2041
FT /*tag= a
FT /product= "Ebola-Zaire virus envelope glycoprotein"
XX
PN WO200008131-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017702.
XX
PR 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
```

```
XX (PURD ) PURDUE RES FOUND.
XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
XX Fischbach MA;
XX WPI; 2000-224030/19.
XX P-PSDB; AAY70074.
XX Cells that produce inventive pseudotyped retroviruses having a broad host
XX range useful for introducing nucleotide sequences into target cells.
XX Example 8; Page 58-60; 65pp; English.
XX
CC The present sequence encodes the Ebola-Zaire virus, subtype Mayinga
CC strain, envelope glycoprotein. Plasmid pcDNA3 was produced by cloning
CC into the polylinker of plasmid pcDNA3 nucleotide sequences corresponding
CC to nucleotides 6029-8253 from the Ebola Zaire virus genome. gpGFP (green
CC fluorescent protein) cells were transiently transfected with pZGFP1. This
CC is used for generation of stable cell lines transiently producing Ebola-
CC MMLV (Moloney murine leukaemia virus) pseudotyped retrovirus including
CC nucleotide sequences encoding GFP in its genome, with a broad host range.
CC These cells having different viral glycoproteins in its lipid bilayer
CC with a pseudotyped retrovirus is transduced with a desired ribonucleo-
CC side sequence. This is used to identify screening agents effective in blocking
CC viral entry into a cell. These agents may be immunological agents like
CC monoclonal or polyclonal antibodies. The pharmacological agents include
CC proteins, peptides or various chemical agents. The pseudotyped retrovirus
CC may be useful in methods of identifying cell surface receptors that allow
CC viral entry
XX
SQ Sequence 2224 BP; 687 A; 541 C; 490 G; 506 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 3; Length 2224;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATATTGCACTTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 60
Db 11 ATGGGCGTTACAGGAATATTGCACTTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 70
Qy 61 CTTTGGGTATATTCCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCCACCAAT 120
Db 71 CTTTGGGTATATTCCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCCACCAAT 130
Qy 121 AGCATTACAGGTTAGTGTGCGACAAACTAGTGTGCTGACAAACTGTCTATCCACA 180
Db 131 AGCATTACAGGTTAGTGTGCGACAAACTAGTGTGCTGACAAACTGTCTATCCACA 190
Qy 181 AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGGAATGGCAACTGAGCTGCCA 240
Db 191 AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGGAATGGCAACTGAGCTGCCA 250
Qy 241 TCTGCACTAAAGATGGGCTTCAGTCCGGTTCCTCCCAAGAGGTTGTTCAATATGAA 300
Db 251 TCTGCACTAAAGATGGGCTTCAGTCCGGTTCCTCCCAAGAGGTTGTTCAATATGAA 310
Qy 301 GCTGTGAATGGGCTGAAACTGCTACAACTCTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 311 GCTGTGAATGGGCTGAAACTGCTACAACTCTGAAATCAAAAAACCTGACGGGAGTGAG 370
Qy 361 TGTCTACAGCAGCCAGACGGGATTCGGGGCTTCCCGGTCCTCCCGGTATGTGCACAAA 420
Db 371 TGTCTACAGCAGCCAGACGGGATTCGGGGCTTCCCGGTCCTCCCGGTATGTGCACAAA 430
Qy 421 GTATCAGGAACGGGACCGGTGCGCGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 480
Db 431 GTATCAGGAACGGGACCGGTGCGCGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 490
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 491 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 550
```

```
QY 541 GTTGCAATTCCTGATCTACTGCCCCAAGCTAAGAGGACTCTTCTCAGCTCAACCCCTTGAGA 600
Db 551 GTTGCAATTCCTGATCTACTGCCCCAAGCTAAGAGGACTCTTCTCAGCTCAACCCCTTGAGA 610
QY 601 GAGCGGTCAATGCACCGGAGGACCGCTAGTGGCTACTATTCTACCAATTAGATAT 660
Db 611 GAGCGGTCAATGCACCGGAGGACCGCTAGTGGCTACTATTCTACCAATTAGATAT 670
QY 661 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGACAAATTTGACC 720
Db 671 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGACAAATTTGACC 730
QY 721 TACGTCCAACTGAATCAAGATTACACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 731 TACGTCCAACTGAATCAAGATTACACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 790
QY 781 TATACAGTGGGAAAGGAGCAATACCGGGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
Db 791 TATACAGTGGGAAAGGAGCAATACCGGGGAAACTAAATTTGGAAGGTCAACCCCGAA 850
QY 841 ATTGATACAAATCGGGAGTGGGCTTCTCGGAAACTTAAAAAAAACCCTCACTAGAAAA 900
Db 851 ATTGATACAAATCGGGAGTGGGCTTCTCGGAAACTTAAAAAAAACCCTCACTAGAAAA 910
QY 901 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAACATCAGTGGT 960
Db 911 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAACATCAGTGGT 970
QY 961 CAGAGTCCGGCGGACACTTCTCCGACCCAGGAGCCACACACAACTGAGACACACAAA 1020
Db 971 CAGAGTCCGGCGGACACTTCTCCGACCCAGGAGCCACACACAACTGAGACACACAAA 1030
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 1031 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1090
QY 1081 GCAGTGTCCGATCTAACAAACCTTGCACAAATCTCCACGAGTCCCAATTCCTCAACACC 1140
Db 1091 GCAGTGTCCGATCTAACAAACCTTGCACAAATCTCCACGAGTCCCAATTCCTCAACACC 1150
QY 1141 AAACCAAGTCCGGACACAGACCCCTATATACACCGGTGTATTAACCTTGACATCTCTGAG 1200
Db 1151 AAACCAAGTCCGGACACAGACCCCTATATACACCGGTGTATTAACCTTGACATCTCTGAG 1210
QY 1201 GCAACTCAAGTTGAACAAATCACCGCAGACAGCAACGACAGCAGCAGCCTCCGACACT 1260
Db 1211 GCAACTCAAGTTGAACAAATCACCGCAGACAGCAACGACAGCAGCAGCCTCCGACACT 1270
QY 1261 CCCTCTGCACGACCGACCGGACCCCAAGAGAGAGAGTCCCAACACGAGGAGCCCTGGC 1320
Db 1271 CCCTCTGCACGACCGACCGGACCCCAAGAGAGAGAGTCCCAACACGAGGAGCCCTGGC 1330
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAACACGAGGAGGAGGAGTCCGAGTCCG 1380
Db 1331 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAACACGAGGAGGAGGAGTCCGAGTCCG 1390
QY 1381 AACCAACACATCATCAACAGATACCGGAGAGAGTCCCAACACGAGGAGGAGTCCGAGTCCG 1440
Db 1391 AACCAACACATCATCAACAGATACCGGAGAGAGTCCCAACACGAGGAGGAGTCCGAGTCCG 1450
QY 1441 TTAATTTACCAATCAATTTCTGGAGTCCGAGGAGTCAATCAGAGCGGGGAGAGAACTCGA 1500
Db 1451 TTAATTTACCAATCAATTTCTGGAGTCCGAGGAGTCAATCAGAGCGGGGAGAGAACTCGA 1510
QY 1501 CGA 1503
Db 1511 AGA 1513
```

RESULT 5
AAZ87189
ID AAZ87189 standard; cDNA; 2298 BP.
XX

AC AAZ87189;
XX 15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX Ebola virus glycoprotein (GP) cDNA.
XX Ebola virus; filovirus; glycoprotein; GP; antigen delivery;
KW genetic vaccine; VEE virus replicon; Venezuelan equine encephalitis;
KW Ebola haemorrhagic fever; protective immunity; diagnosis; ss.
XX Ebola virus; 'Zaire 1976'.
OS
XX
XX Key Location/Qualifiers
FT CDS 59..2089
FT /*tag= a
FT /product= "Ebola virus glycoprotein (GP)"
XX WO200000617-A2.
XX
XX 06-JAN-2000.
XX 22-JUN-1999; 99WO-US014311.
XX 29-JUN-1998; 98US-0091403P.
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX Hart MK, Wilson JA, Pushko P, Smith JP, Schmaljohn AL;
PI WPI; 2000-160677/14.
XX P-PSDB; AAY77112.
XX
XX New GP, NP, VP24, VP30, VP35 and VP40 Ebola virus proteins, useful for
PT prevention, treatment or diagnosis of Ebola infection, particularly where
PT expressed from virus replicons.
XX
XX Claim 1; Page 61; 71pp; English.
XX The invention relates to methods and compositions for inducing a
CC protective immune response against Ebola virus. The Ebola virus (Zaire
CC 1976, Mayinga isolate) genes encoding glycoprotein (GP), nucleoprotein
CC (NP), and structural proteins VP24, VP30, VP30#2, VP35 and VP40 were
CC individually inserted into a Venezuelan equine encephalitis virus (VEE)
CC replicon vector. The Ebola virus gene replaces the VEE virus structural
CC protein genes; the result is a self-replicating RNA molecule that encodes
CC its own replicase and transcriptase functions, and in addition makes
CC abundant quantities of the Ebola virus protein. When replicon RNA is
CC transfected into eukaryotic cells, along with two helper RNAs that
CC express the VEE virus structural proteins, the replicon RNA is packaged
CC into VEE virus-like particles by the VEE virus structural proteins, which
CC are provided in trans. Since the helper RNAs lack packaging signals
CC necessary for further propagation, the resulting VEE replicon particles
CC (VRPs) which are produced are infectious for one cycle but are defective
CC whereby the cell produces the Ebola virus protein, is ultimately killed
CC by the infection, but does not produce any viral progeny. VRPs containing
CC a gene encoding an Ebola virus protein were injected subcutaneously into
CC mice. VRPs expressing Ebola GP, NP, VP24, VP30, VP30#2, VP35 or VP40
CC genes induced protective immunity in mice and may provide protection in
CC humans. The Ebola virus proteins, nucleic acids, replicons, vectors and
CC constructs may be used to produce vaccines against Ebola virus infection
CC (Ebola haemorrhagic fever) in mammals, to elicit immune responses against
CC Ebola antigens, to confer protective immunity, and to reduce disease
CC symptoms and reduce the severity of disease. Ebola proteins and specific
CC antibodies may be used for the diagnosis of Ebola infection. DNA
CC sequences encoding the Ebola virus proteins may be useful as probes and
CC primers for diagnostic hybridisation or PCR assays for detecting Ebola
CC virus. The present sequence represents a cDNA encoding Ebola virus
CC glycoprotein (GP). (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;

Query Match		66.7%;	Score 1501.4;	DB 3;	Length 2298;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 1502;		Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
QY	1	ATGGCGGTTACAGGAATATTG	CAGTATTG	CAGTATTG	CAGTATTG	CAGTATTG
DB	59	ATGGCGGTTACAGGAATATTG	CAGTATTG	CAGTATTG	CAGTATTG	CAGTATTG
QY	61	CTTTGGGTAATTATCTTTT	CAAGAAAT	TTTCCAT	CCCACT	TTGGAGTATCCACAAT
DB	119	CTTTGGGTAATTATCTTTT	CAAGAAAT	TTTCCAT	CCCACT	TTGGAGTATCCACAAT
QY	121	AGCACATTACAGGTTAGT	GTGACAAAT	PAGTTT	GTGACAAAT	PAGTTT
DB	179	AGCACATTACAGGTTAGT	GTGACAAAT	PAGTTT	GTGACAAAT	PAGTTT
QY	181	AATCAATTGAGATCAGTT	GGATCT	CGAAGGGAAT	TGAGTGGCAACT	GACGTGCCA
DB	239	AATCAATTGAGATCAGTT	GGATCT	CGAAGGGAAT	TGAGTGGCAACT	GACGTGCCA
QY	241	TCTGCAACTAAAGAT	GGGCTT	CAGTCCG	GTGTC	CAACAAAGGTGGTCAATATGAA
DB	299	TCTGCAACTAAAGAT	GGGCTT	CAGTCCG	GTGTC	CAACAAAGGTGGTCAATATGAA
QY	301	GCTGTGTAATGGCTG	AAATCT	CTTGAAT	CAAAAT	CTTGAATCAAAATCTGACGGAGTGCAG
DB	359	GCTGTGTAATGGCTG	AAATCT	CTTGAAT	CAAAAT	CTTGAATCAAAATCTGACGGAGTGCAG
QY	361	TGTTCTACAGCAGCG	CCAGAC	CGGGAT	TCGGGCT	TTCCCGGTCGGGTATGTGCACAAA
DB	419	TGTTCTACAGCAGCG	CCAGAC	CGGGAT	TCGGGCT	TTCCCGGTCGGGTATGTGCACAAA
QY	421	GTATCAGGAA	CGGAC	CGGTG	CCGAGAC	TTTTCCTTCCATAAAGAGGGTCTTTCTTC
DB	479	GTATCAGGAA	CGGAC	CGGTG	CCGAGAC	TTTTCCTTCCATAAAGAGGGTCTTTCTTC
QY	481	CTGTATGATCGACT	TTGCTT	CCACAGT	TATCTAC	CGAGGAACGACTTTTCGCTGAAGGTGTC
DB	539	CTGTATGATCGACT	TTGCTT	CCACAGT	TATCTAC	CGAGGAACGACTTTTCGCTGAAGGTGTC
QY	541	GTTCATTTCTGAT	ACTTGC	CCCAAGT	AAAGAGT	ACTTTTCAGCTCACACCCCTTGAGA
DB	599	GTTCATTTCTGAT	ACTTGC	CCCAAGT	AAAGAGT	ACTTTTCAGCTCACACCCCTTGAGA
QY	601	GAGCCGGTCAATG	CAACCG	AGGAC	CCGCT	CTAGTGGCTACTATTCTACCAATATAGATAT
DB	659	GAGCCGGTCAATG	CAACCG	AGGAC	CCGCT	CTAGTGGCTACTATTCTACCAATATAGATAT
QY	661	CAGGCTACCGG	TTTTGG	AAACCA	TAGACAG	AGTACTTTGTCGAGTTGACAAATTTGACC
DB	719	CAGGCTACCGG	TTTTGG	AAACCA	TAGACAG	AGTACTTTGTCGAGTTGACAAATTTGACC
QY	721	TACGTCCAACTT	GAAATCA	AGATT	TCAC	CACAGTTTCTGCTCCAGCTCAATGAGACAATA
DB	779	TACGTCCAACTT	GAAATCA	AGATT	TCAC	CACAGTTTCTGCTCCAGCTCAATGAGACAATA
QY	781	TATACAAGTGG	AAAGAG	CAATACC	CGGAA	AACTAATTTGGAAGGTCAACCCCGAA
DB	839	TATACAAGTGG	AAAGAG	CAATACC	CGGAA	AACTAATTTGGAAGGTCAACCCCGAA
QY	841	ATTGATACAA	CAATCG	GGAGTGG	GGCTT	CTGGGAACTAATAAAAACTCTACTAGAAAA
DB	899	ATTGATACAA	CAATCG	GGAGTGG	GGCTT	CTGGGAACTAATAAAAACTCTACTAGAAAA
QY	901	ATTTCAGTGA	AGAGTGT	CTTTT	CAAGT	TTGATCAACCGGACCAAAAAATCATCAGTGT
DB	959	ATTTCAGTGA	AGAGTGT	CTTTT	CAAGT	TTGATCAACCGGACCAAAAAATCATCAGTGT
QY	961	CAGAGTCCG	CGCGAA	CTTTCT	CCGACC	CGGAGCAACACAACTCTGAAGACCAAAA
DB	1019	CAGAGTCCG	CGCGAA	CTTTCT	CCGACC	CGGAGCAACACAACTCTGAAGACCAAAA
QY	1021	ATCATGGCTT	CAGAAAA	TTCTCT	CGAAT	GGTTCAAGTGCACAGTCAAGGAAGGAAGCT

The present sequence encodes Ebola virus GP, a type I transmembrane

CC protein. The invention describes novel Ebola GP monoclonal antibodies and
CC epitopes recognised by these antibodies. Individual antibodies or
CC mixtures of antibodies are useful for preventing Ebola virus infection in
CC a subject. They are useful for inhibiting and ameliorating symptoms of
CC Ebola virus infection. They are also useful for detecting Ebola virus in
CC a sample, by incubating the sample with the antibodies, and detecting
CC antibody-Ebola virus complexes. The antibodies are also useful for
CC treating Ebola virus infection

XX Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 4; Length 2298;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCGGTACAGGAATATTCAGTTACCTCGTATCGATTCAAGAGGACATCATCTTT 60
DB 59 ATGGCGGTACAGGAATATTCAGTTACCTCGTATCGATTCAAGAGGACATCATCTTT 118
QY 61 CTTTGGGTAATATTCCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCCAAT 120
DB 119 CTTTGGGTAATATTCCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCCAAT 178
QY 121 AGCATTACAGGTTAGTATTCGACAAATAGTTTGTGTCGACAAATCTGTCATCCACA 180
DB 179 AGCATTACAGGTTAGTATTCGACAAATAGTTTGTGTCGACAAATCTGTCATCCACA 238
QY 181 ATCAATTCAGATCAGTTGAGTATTCGAGGATTCGAGGATTCGAGGATTCGAGGATTC 240
DB 239 ATCAATTCAGATCAGTTGAGTATTCGAGGATTCGAGGATTCGAGGATTCGAGGATTC 298
QY 241 TCTGCACTAAAGATATGGGCTTTCAGGTCGGGTCCCAACCAAGGTGCTCAATATGAA 300
DB 299 TCTGCACTAAAGATATGGGCTTTCAGGTCGGGTCCCAACCAAGGTGCTCAATATGAA 358
QY 301 GCTGTGAATGGCTGAAATCTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
DB 359 GCTGTGAATGGCTGAAATCTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 418
QY 361 TGTCTACGACGACCGGATTCGGGCTTCCCGGCTTCCCGGCTGATGTCGACAAA 420
DB 419 TGTCTACGACGACCGGATTCGGGCTTCCCGGCTTCCCGGCTGATGTCGACAAA 478
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGCTTCTTCTTC 480
DB 479 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGCTTCTTCTTC 538
QY 481 CTGTATGATCGATCTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 540
DB 539 CTGTATGATCGATCTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 598
QY 541 GTTGCATTTCTGATCTGCTTCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGGA 600
DB 599 GTTGCATTTCTGATCTGCTTCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGGA 658
QY 601 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB 659 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCAACAATTAGATAT 718
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTGACC 720
DB 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTGACC 778
QY 721 TAGGTCACATTTGAATCAAGATTCACACAGTTTCTGCTCAGCTGATGAGACATAT 780
DB 779 TAGGTCACATTTGAATCAAGATTCACACAGTTTCTGCTCAGCTGATGAGACATAT 838
QY 781 TATACAGTGGGAAAGGAGCAATACACGGAATAAATTTGGAAGGTCAACCCCGAA 840
DB 839 TATACAGTGGGAAAGGAGCAATACACGGAATAAATTTGGAAGGTCAACCCCGAA 898
QY 841 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAAACTTAAAAAAAACCCTCATGAAAA 900
```

```
DB 899 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAAACTAAAAAAACCTCACTAGAAA 958
QY 901 ATTCGCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACGGAGCAAAAAACATCAGTGT 960
DB 959 ATTCGCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACGGAGCAAAAAACATCAGTGT 1018
QY 961 CAGAGTCCGGCGGCAACTTCTTTCGAGCCAGGAGCAACACAACTGAAGACCAAAA 1020
DB 1019 CAGAGTCCGGCGGCAACTTCTTTCGAGCCAGGAGCAACACAACTGAAGACCAAAA 1078
QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
DB 1079 ATCATGGCTTCAGAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1138
QY 1081 GCAGTGTGCGCATCTAAACAACCTTTGCCACAACTCTCCAGAGTCCCAATCCCTCAACACC 1140
DB 1139 GCAGTGTGCGCATCTAAACAACCTTTGCCACAACTCTCCAGAGTCCCAATCCCTCAACACC 1198
QY 1141 AAACAGGTCCGGACAAACAGCACCCCATTAATACACCCGTGTATATACTTGACATCTCTGAG 1200
DB 1199 AAACAGGTCCGGACAAACAGCACCCCATTAATACACCCGTGTATATACTTGACATCTCTGAG 1258
QY 1201 GCAACTCAAGTTGAACACATCACCGCAGACACAGACAGACAGACAGACAGACAGACAG 1260
DB 1259 GCAACTCAAGTTGAACACATCACCGCAGACACAGACAGACAGACAGACAGACAGACAG 1318
QY 1261 CCCTCTGCACGACCGCAGCGGACCCCAAAAGCAGAGAAACACCAACAGAGCAAGAGC 1320
DB 1319 CCCTCTGCACGACCGCAGCGGACCCCAAAAGCAGAGAAACACCAACAGAGCAAGAGC 1378
QY 1321 ACTGACTTCTGGACCCCGCACCAACAAAGTCCCAAAAACCAAGAGCAAGACCCCTGGC 1380
DB 1379 ACTGACTTCTGGACCCCGCACCAACAAAGTCCCAAAAACCAAGAGCAAGACCCCTGGC 1438
QY 1381 AACACACACTCATCCAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTTAGGC 1440
DB 1439 AACACACACTCATCCAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTTAGGC 1498
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCAGAGCGGGAGAGAACTCGA 1500
DB 1499 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCAGAGCGGGAGAGAACTCGA 1558
QY 1501 CGA 1503
DB 1559 AGA 1561
```

RESULT 7

ADL27460

ID ADL27460 standard; DNA; 2298 BP.

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

XX AC ADL27460;

PR 23-AUG-2002; 2002US-00226795.
XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
PA Hart MK, Wilson JA;
PI WPI; 2004-226835/21.
DR P-PSDB; ADL27461.
DR
XX New isolated monoclonal antibody that binds Ebola virus GP, which
PT monoclonal antibody comprises a heavy chain variable region, useful for
PT treating or ameliorating Ebola virus infection.
XX
PS Disclosure; Page 58-59; 68pp; English.
XX
CC The specification describes monoclonal antibodies that bind Ebola virus
CC glycoprotein. The antibodies are EGP 6D8-1-1, EGP 13F6-1-2, and EGP 13C6-
CC 1-1. The monoclonal antibodies of the invention are useful for treating
CC or ameliorating an Ebola virus infection. The antibodies were found to be
CC effective when administered 2 days after challenge, after significant
CC viral replication had occurred. The present sequence encodes an Ebola
CC glycoprotein. Monoclonal antibodies of the invention recognise epitopes
CC from this protein.
XX
SQ Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 12; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTGAGTTACCTCGTGATGATTCAGAGGACATCTCTTT 60
DB |||||||
QY 59 ATGGGCGTTACAGGAATATTGAGTTACCTCGTGATGATTCAGAGGACATCTCTTT 118
DB |||||||
QY 61 CTTTGGGTAATATTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
DB |||||||
QY 119 CTTTGGGTAATATTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 178
DB |||||||
QY 121 AGCATTATACAGGTAGTGTGACAACTAGTTTGTGCTGATGACAACTGTGATCCACA 180
DB |||||||
QY 179 AGCATTATACAGGTAGTGTGACAACTAGTTTGTGCTGATGACAACTGTGATCCACA 238
DB |||||||
QY 181 AATCAATTTGAGATCAGTTGGATGATCTCGAAGGGAATGAGTGGCACTGAGTGCCA 240
DB |||||||
QY 239 AATCAATTTGAGATCAGTTGGATGATCTCGAAGGGAATGAGTGGCACTGAGTGCCA 298
DB |||||||
QY 241 TCTGCAACTTAAAGATGGGCTTCAAGTCCGGTGTCCCAACAAAGGTTGGTCAATTTAGAA 300
DB |||||||
QY 299 TCTGCAACTTAAAGATGGGCTTCAAGTCCGGTGTCCCAACAAAGGTTGGTCAATTTAGAA 358
DB |||||||
QY 301 GCTGTGAAATGGGCTGAAATCTGTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 360
DB |||||||
QY 359 GCTGTGAAATGGGCTGAAATCTGTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 418
DB |||||||
QY 361 TGTCTACAGCAGCCGACGAGATTCGGGCTTCCCGGCTTCCCGGCTGATGTGACAAA 420
DB |||||||
QY 419 TGTCTACAGCAGCCGACGAGATTCGGGCTTCCCGGCTTCCCGGCTGATGTGACAAA 478
DB |||||||
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 480
DB |||||||
QY 479 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 538
DB |||||||
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
DB |||||||
QY 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 598
DB |||||||
QY 541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTCAGCTCACACCCCTTGAGA 600
DB |||||||
QY 599 GTTGCATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTTTCAGCTCACACCCCTTGAGA 658
DB |||||||
QY 601 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT 660
DB |||||||
QY 659 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT 718
DB |||||||

RESULT 8
ACC71526
ID ACC71526 standard; DNA; 6887 BP.
XX ACC71526;
XX
DT 10-JUL-2003 (first entry)
XX
DS VRC6005 (pVR1012-GP(Z) delta GP2 delta C-term A) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;

QY 661 CAGGCTACCGGTTTTTGAAACAATGACACAGAGTACTTGTTCGAGGTTGACAAATTGACC 720
DB |||||||
QY 719 CAGGCTACCGGTTTTTGAAACAATGACACAGAGTACTTGTTCGAGGTTGACAAATTGACC 778
DB |||||||
QY 721 TACGTCCAACTTGAATCAAGATTACACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 780
DB |||||||
QY 779 TACGTCCAACTTGAATCAAGATTACACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 838
DB |||||||
QY 781 TATCAAGTGGGAAAAGAGCAATATACACGGGAAAACCTAAATTTGGAAGGTCAACCCCGAA 840
DB |||||||
QY 839 TATCAAGTGGGAAAAGAGCAATATACACGGGAAAACCTAAATTTGGAAGGTCAACCCCGAA 898
DB |||||||
QY 841 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAAACTAATAAATAACCTCCTAGAAAA 900
DB |||||||
QY 899 ATTGATACAACTATCGGGAGTGGGCTTCTGGGAAACTAATAAATAACCTCCTAGAAAA 958
DB |||||||
QY 901 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAAGGAGCAAAAACATCAGTGGT 960
DB |||||||
QY 959 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAAGGAGCAAAAACATCAGTGGT 1018
DB |||||||
QY 961 CAGAGTCCGGCGGAACTTTCTCCGACCCAGGGACCAACACAACTGAAGACCAACAA 1020
DB |||||||
QY 1019 CAGAGTCCGGCGGAACTTTCTCCGACCCAGGGACCAACACAACTGAAGACCAACAA 1078
DB |||||||
QY 1021 ATCATGCTTCAGAAAATTTCTCTGCAATGTTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB |||||||
QY 1079 ATCATGCTTCAGAAAATTTCTCTGCAATGTTTCAAGTGCACAGTCAAGGAAGGAAGCT 1138
DB |||||||
QY 1081 GCAGTGTGCAATCTAAACAACTTTGCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
DB |||||||
QY 1139 GCAGTGTGCAATCTAAACAACTTTGCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1198
DB |||||||
QY 1141 AAACAGGTCGGGACAAACAGCACCCCAATAATACACCCGTGTATAAACTTTGACATCTCTGAG 1200
DB |||||||
QY 1199 AAACAGGTCGGGACAAACAGCACCCCAATAATACACCCGTGTATAAACTTTGACATCTCTGAG 1258
DB |||||||
QY 1201 GCAACTCAAGTTGAACAACATATCAGGAGAAACAGACAAACAGACACAGCCCTCGACACT 1260
DB |||||||
QY 1259 GCAACTCAAGTTGAACAACATATCAGGAGAAACAGACAAACAGACACAGCCCTCGACACT 1318
DB |||||||
QY 1261 CCCTCTCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACACAGGCAAGAGC 1320
DB |||||||
QY 1319 CCCTCTCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACACAGGCAAGAGC 1378
DB |||||||
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGAGACCGCTGGC 1380
DB |||||||
QY 1379 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGAGACCGCTGGC 1438
DB |||||||
QY 1381 AACAAACACTCATCAACAAAGATACCGGAGAAAGAGTGCAGAGCGGGAAGCTAGGC 1440
DB |||||||
QY 1439 AACAAACACTCATCAACAAAGATACCGGAGAAAGAGTGCAGAGCGGGAAGCTAGGC 1498
DB |||||||
QY 1441 TTAATTAACCAATATCTTGTGGAGTGCAGGACTGTATCAGGCGGGAGAAAGCTCGA 1500
DB |||||||
QY 1499 TTAATTAACCAATATCTTGTGGAGTGCAGGACTGTATCAGGCGGGAGAAAGCTCGA 1558
DB |||||||
QY 1501 CGA 1503
DB |||
QY 1559 AGA 1561
DB |||

Db 3379 TTAATTACCAATACTATTGCTGGAGTCGACGAGTGTATCATCAGCGGGAGAGACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
RESUFT 9
ACC71529
ID ACC71529 standard; DNA; 6914 BP.
XX
AC ACC71529;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6008 (pVR1012-GP(Z) delta TM) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
PN W02003028632-A2. -
XX
PD 10-APR-2003.
XX
PP 24-SEP-2002; 2002WO-US030251.
XX
PR 01-OCT-2001; 2001US-0326476P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
DR WPI; 2003-371961/35.
XX
PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 133-135; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprising of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprising of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC truncation of the TM region of Ebola Glycoprotein of Zaire Subtype
CC (GP(Z))
XX
SQ Sequence 6914 BP; 1850 A; 1750 C; 1645 G; 1669 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 6914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGGTTACAGGAATATTCAGTTACCTCTGATTCGATTCAAGAGGACATCATCTTTT 60
Db 1939 ATGGGGGTTACAGGAATATTCAGTTACCTCTGATTCGATTCAAGAGGACATCATCTTT 1998
QY 61 CTTTGGGTAATTCCTTTTCCAGAGACATTTCCATCCCACTTGGAGTCCACCAAT 120
|||||

Db 1999 CTTTGGGTAATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCCACCAAT 2058
QY 121 AGCACATTACAGGTTAGTGTGCGACAAACTAGTTGTCTGTGACAAACTGTCTCATCCACA 180
Db 2059 AGCACATTACAGGTTAGTGTGCGACAAACTAGTTGTCTGTGACAAACTGTCTCATCCACA 2118
QY 181 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCGCA 240
Db 2119 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCGCA 2178
QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACAAAGGTGCTCAATTTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACAAAGGTGCTCAATTTATGAA 2238
QY 301 GCTGGTGAATGGGCTGAAACTGCTCAATCTTGAATCAAAAAAACCCTGACGGGAGTGAG 360
Db 2239 GCTGGTGAATGGGCTGAAACTGCTCAATCTTGAATCAAAAAAACCCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAAACGACTTTTCGCTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAAACGACTTTTCGCTGAAGGTGTC 2478
QY 541 GTTGCATTTCTGATCTACGCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATCTACGCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGGGAGACCCGCTCTAGTGGCTACTATTCTACCACAATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGGGAGACCCGCTCTAGTGGCTACTATTCTACCACAATTAGATAT 2598
QY 661 CAGGTACCGGTTTGGAAACCAATGAGACAGACTTCTGTTGCGAGTTGACAAATTTGACC 720
Db 2599 CAGGTACCGGTTTGGAAACCAATGAGACAGACTTCTGTTGCGAGTTGACAAATTTGACC 2658
QY 721 TAGCTCCAATCTGAATCAAGATTACACACAGTTTCTGCTCCAGCTCAATGAGACATA 780
Db 2659 TAGCTCCAATCTGAATCAAGATTACACACAGTTTCTGCTCCAGCTCAATGAGACATA 2718
QY 781 TATCAAGTGGGAAAGGAGCAATACACACGGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATCAAGTGGGAAAGGAGCAATACACACGGGAAACTAAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAAATCGGGGAGTGGGCTTTCTGGGAAACTAAAAAAAACCCTCACTAGAAAA 900
Db 2779 ATTGATACAAATCGGGGAGTGGGCTTTCTGGGAAACTAAAAAAAACCCTCACTAGAAAA 2838
QY 901 ATTGCGAGTGAAGATTGTCTTTTACAGTTCTATCAACGGGAGCCCAAAACATCAGTGGT 960
Db 2839 ATTGCGAGTGAAGATTGTCTTTTACAGTTGTATCAACGGGAGCCCAAAACATCAGTGGT 2898
QY 961 CAGAGTCCGGCGCGAACTTTCTTCGACCCAGGAGCCAAACAACTGAAGAGCCACAAA 1020
Db 2899 CAGAGTCCGGCGCGAACTTTCTTCGACCCAGGAGCCAAACAACTGAAGAGCCACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAAAGCT 3018
QY 1081 GCAGTGTGCAATCAAGAAACCTTTGGCCCAATCTCCACAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCAATCAAGAAACCTTTGGCCCAATCTCCACAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACCAGGTCCGGCAACAGCACCCATATAACCCCGTGTATAAATTTGACATCTCTGAG 1200
Db 3079 AAACCAGGTCCGGCAACAGCACCCATATAACCCCGTGTATAAATTTGACATCTCTGAG 3138


```
Db 2839 ATTCCAGTGAAGATTGCTTTTCAGTTGATCAACGGAGCCAAAACATCATGTGT 2898
Qy 961 CAGAGTCCGGCGGAATCTTTCCGACCCAGGAGCAACACAACTGAAGACACAAA 1020
Db 2899 CAGAGTCCGGCGGAATCTTTCCGACCCAGGAGCAACACAACTGAAGACACAAA 2958
Qy 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAGGAAGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAGGAAGGAAGCT 3018
Qy 1081 GCAGTGTGCATCTAACAACTTCGCCAATCTCCACAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCATCTAACAACTTCGCCAATCTCCACAGTCCCAATCCCTCAACACC 3078
Qy 1141 AAACAGGTCCGGCAACAGCACCCATATACACCCGTGTATATAACTTGCATCTCTGAG 1200
Db 3079 AAACAGGTCCGGCAACAGCACCCATATACACCCGTGTATATAACTTGCATCTCTGAG 3138
Qy 1201 GCAACTCAAGTTGAACACATCACCGCAGACAGACAGACAGACAGACAGACAGAC 1260
Db 3139 GCAACTCAAGTTGAACACATCACCGCAGACAGACAGACAGACAGACAGACAGAC 3198
Qy 1261 CCTCTGCCACGACCGCGAGCCGCCCAAGAGCAGAGAACCAACACAGCAGCAAGAGC 1320
Db 3199 CCTCTGCCACGACCGCGAGCCGCCCAAGAGCAGAGAACCAACACAGCAGCAAGAGC 3258
Qy 1321 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCCCTGGC 1380
Db 3259 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCCCTGGC 3318
Qy 1381 AACACACACATCATCAACAGATACCGGAGAGAGAGTGCCAGACGGGGAAGCTAGGC 1440
Db 3319 AACACACACATCATCAACAGATACCGGAGAGAGAGTGCCAGACGGGGAAGCTAGGC 3378
Qy 1441 TTAATTTACCAATATTGCTGAGTCCGAGCTGATCAGACGGCGGAGAGACTCGA 1500
Db 3379 TTAATTTACCAATATTGCTGAGTCCGAGCTGATCAGACGGCGGAGAGACTCGA 3438
Qy 1501 CGA 1503
Db 3439 AGA 3441

RESULT 11
ACC71528
ID ACC71528 standard; DNA; 7106 BP.
AC ACC71528;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6007 (pVR1012-OP(Z) delta GP2 delta FUS) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
XX viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
PN WO2003028632-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030251.
XX
PR 01-OCT-2001; 2001US-0326476P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX WPI; 2003-371961/35.
DR
```

```
XX
PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 130-132; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC a deletion of a fusion peptide in GP2 of Ebola Glycoprotein of Zaire
CC Subtype (GP(Z))
XX
SQ Sequence 7106 BP; 1923 A; 1770 C; 1667 G; 1746 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 10; Length 7106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGTTACAGGAATATTCAGTTACCTCGTATCGATTCAAGAGACATCATCTTT 60
Db 1939 ATGGGCGTTACAGGAATATTCAGTTACCTCGTATCGATTCAAGAGACATCATCTTT 1998
Qy 61 CTTTGGGTAAATATCTCTTCCAAAGACATTTTCCATCCCATCTGGAGTCATCCAAAT 120
Db 1999 CTTTGGGTAAATATCTCTTCCAAAGACATTTTCCATCCCATCTGGAGTCATCCAAAT 2058
Qy 121 AGCATTATACAGGTAGTGTGACAAATAGTCTGTCGACAAATCTGTCATCCACA 180
Db 2059 AGCATTATACAGGTAGTGTGACAAATAGTCTGTCGACAAATCTGTCATCCACA 2118
Qy 181 AATCAATGAGATCAGTTGGAATCTCGAAGGAATGAGTGGCACTGACCTGCCA 240
Db 2119 AATCAATGAGATCAGTTGGAATCTCGAAGGAATGAGTGGCACTGACCTGCCA 2178
Qy 241 TCTGCNACTAAAGATGGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 300
Db 2179 TCTGCNACTAAAGATGGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 2238
Qy 301 GCTGTGATGGGTGAAATCTGCTACAAATCTGAAATCAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGATGGGTGAAATCTGCTACAAATCTGAAATCAAAAACCTGACGGGAGTGAG 2298
Qy 361 TGTCTACAGACGCCACAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGACGCCACAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
Qy 421 GTATCAGGAACGGGACCGGTGCGGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGCGGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 2418
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACAGGAAACGACTTTCCGTGAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACAGGAAACGACTTTCCGTGAGGTGTC 2478
Qy 541 GTTGCAATTTCTGATCTGCCCCCAAGCTAAGAGGACTTTCTTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATCTGCCCCCAAGCTAAGAGGACTTTCTTTCAGCTCAACCCCTTGAGA 2538
Qy 601 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGTACTATTCTTACCACAAATTAGATAT 660
Db 601 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGTACTATTCTTACCACAAATTAGATAT 660
```



```
Db 2539 GAGCCGGTCAATGCAACGAGGACCGTCTAGTGGCTACTATTCTACCAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTCTGTCAGGTTGACAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTCTGTCAGGTTGACAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTTCAACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTTCAACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAGTGGGAAAGAGGCAATACCAACGGGAAACTAATTTGGAGTCAACCCCGAA 840
Db 2719 TATACAGTGGGAAAGAGGCAATACCAACGGGAAACTAATTTGGAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGCGAGTGGGCTTCTCGGAACTTAAATAAACTCTACTAGAAAA 900
Db 2779 ATTGATACAACTCGGCGAGTGGGCTTCTCGGAACTTAAATAAACTCTACTAGAAAA 2838
QY 901 ATTCCAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGGACCAAAAAACATCAGTGGT 960
Db 2839 ATTCCAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGGACCAAAAAACATCAGTGGT 2898
QY 961 CAGAGTCCGGCGGGAATCTTCTCGACCCAGGACCAACACAACTGAAGACACACAAA 1020
Db 2899 CAGAGTCCGGCGGGAATCTTCTCGACCCAGGACCAACACAACTGAAGACACACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAGCT 3018
QY 1081 GCAGTGTGCATCTTAACAACCCCTTGCCACAATCTCCACAGTCCCAATCCCTCAACCC 1140
Db 3019 GCAGTGTGCATCTTAACAACCCCTTGCCACAATCTCCACAGTCCCAATCCCTCAACCC 3078
QY 1141 AAACCCAGTCCGACACAGACCCCNATATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 3079 AAACCCAGTCCGACACAGACCCCNATATACACCGGTGTATAAATTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACATACCGCAGAACAGACAGACAGACAGACGCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACATACCGCAGAACAGACAGACAGACAGACGCTCCGACACT 3198
QY 1261 CCCTCTGCCACAGCCGACCGGACCCCAAAAGCAGAGAGAACCAACACGAGCAGAGAGC 1320
Db 3199 CCCTCTGCCACAGCCGACCGGACCCCAAAAGCAGAGAGAACCAACACGAGCAGAGAGC 3258
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACACACAGGAGACCCGCTGGC 1380
Db 3259 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACACACAGGAGACCCGCTGGC 3318
QY 1381 AACACACACATCATCAACCAAGATACCGGAGAGAGAGTCCAGACGCGGAAGCTTAGGC 1440
Db 3319 AACACACACATCATCAACCAAGATACCGGAGAGAGAGTCCAGACGCGGAAGCTTAGGC 3378
QY 1441 TTAATTTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGGCGGGAGAGAACTCGA 1500
Db 3379 TTAATTTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGGCGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
```

RESULT 12

ACCT71521

ID ACC71521 standard; DNA; 7154 BP.

XX AC

AC ACC71521;

XX DT

DT 10-JUL-2003 (first entry)

XX DE

DE VRC6000 (pVR1012-GP(2) plasmid.

XX XX

```
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX WO2003028632-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030251.
XX 01-OCT-2001; 2001US-0326476P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 114-116; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, or
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid expresses the Ebola Glycoprotein of Zaire
CC subtype and is derived from a pVR1012 backbone
XX
SQ Sequence 7154 BP; 1932 A; 1784 C; 1682 G; 1756 T; 0 U; 0 Other;
```

Query Match 66.7%; Score 1501.4; DB 10; Length 7154;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 60

Db 1939 ATGGCGGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 1998

QY 61 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAT 120

Db 1999 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAT 2058

QY 121 AGCACATTACAGGTTAGTGTGACAACTAGTTTTCGTGACAACTGCATCCACA 180

Db 2059 AGCACATTACAGGTTAGTGTGACAACTAGTTTTCGTGACAACTGCATCCACA 2118

QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240

Db 2119 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 2178

QY 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCGCCACCAAGGTGGTCAATTATGAA 300

Db 2179 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCGCCACCAAGGTGGTCAATTATGAA 2238

QY 301 GCTGGTGAATGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGGAGTGAG 360

Db 2239 GCTGGTGAATGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGGAGTGAG 2298

QY 361 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 2478
QY 541 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGAGGAGACCGGTCTAGTGGCTACTATTCTACACACAATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGAGGAGACCGGTCTAGTGGCTACTATTCTACACACAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGTCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGTCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAACTGGGAAAGAGCAATACACACGGGAAACCTAAATTTGGAAGTCAACCCCGAA 840
Db 2719 TATACAACTGGGAAAGAGCAATACACACGGGAAACCTAAATTTGGAAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTTAAATAAACTCACTAGAAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTTAAATAAACTCACTAGAAAAA 2838
QY 901 ATTCCAGTGAAGAGTGTCTTTCACAGTTGTATCAACGGGACCAAAACATCAGTGT 960
Db 2839 ATTCCAGTGAAGAGTGTCTTTCACAGTTGTATCAACGGGACCAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCAACACAACTGMAAGACCACAAA 1020
Db 2899 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCAACACAACTGMAAGACCACAAA 2958
QY 1021 ATCATGGTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 2959 ATCATGGTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 3018
QY 1081 GCAGTGTGCATCTAACCAACCTTGGCAATCTCCAGAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCATCTAACCAACCTTGGCAATCTCCAGAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACAGGTTCGGGACACAGACCCCAATATACACCCGTGTATAAATTTGACATCTCTGAG 1200
Db 3079 AAACAGGTTCGGGACACAGACCCCAATATACACCCGTGTATAAATTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACATCAGCGAGAACAGACAAACAGACAGACAGCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACATCAGCGAGAACAGACAAACAGACAGACAGCTCCGACACT 3198
QY 1261 CCCTCTGCCACACAGCGAGCGGACCCCAAAAGAGAGAGAACACCAACAGGAGGAGAGC 1320
Db 3199 CCCTCTGCCACACAGCGAGCGGACCCCAAAAGAGAGAGAACACCAACAGGAGGAGAGC 3258
QY 1321 ACTGACTTCTCGACCCCGCCACCAACAAAGTCTCCCAAAACACAGCGAGAGCTCTGGC 1380
Db 3259 ACTGACTTCTCGACCCCGCCACCAACAAAGTCTCCCAAAACACAGCGAGAGCTCTGGC 3318
QY 1381 AACAAACACATCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGAGGAGCTAGGC 1440
Db 3319 AACAAACACATCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGAGGAGCTAGGC 3378
QY 1441 TTAATTACCAATACTATTGCTGGAGTGGAGGACTGATCAACAGCGGGGAGAGAACTCGA 1500

Db 3379 TTAATTACCAATACTATTGCTGGAGTGGAGGACTGATCAACAGCGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
RESULT 13
ACC71522
ID ACC71522 standard; DNA; 7188 BP.
XX ACC71522;
AC ACC71522;
XX 10-JUL-2003 (first entry)
XX VRC6001 (pVR1012x/s-GP(Z) plasmid.
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX WO2003028632-A2.
PN 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030251.
XX 01-OCT-2001; 2001US-0326476P.
PR (USSH) US DBPT HEALTH & HUMAN SERVICES.
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
DR New bimodal priming and boosting compositions, useful as viral vaccines,
XX specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
PT Claim 1; Page 116-119; 219pp; English.
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid expresses the Ebola Glycoprotein of Zaire
CC Subtype and is derived from a pVR1012 backbone
XX Sequence 7188 BP; 1942 A; 1791 C; 1689 G; 1766 T; 0 U; 0 Other;
SQ Query Match 66.7%; Score 1501.4; DB 10; Length 7188;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 60
Db 1939 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 1998
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGATCATCAGAT 120

Db 1999 CTTTGGGTAAATATCCTTTTCCAAAGAACATTTTCCATCCCACTTTGGAGTATCCACAAAT 2058
Qy 121 AGCAATTTACAGGTTAGTGATCGACAAATAGTGTTCGTGACAAATAGTGTTCATCCACA 180
Db 2059 AGCAATTTACAGGTTAGTGATCGACAAATAGTGTTCGTGACAAATAGTGTTCATCCACA 2118
Qy 181 AATCAATTTAGATCATGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTTAGATCATGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 2178
Qy 241 TCTGCAATTAAGATGGGCTTCAGTCCGGTGTCCACCAAAAGGTGGTCAATTTATGAA 300
Db 2179 TCTGCAATTAAGATGGGCTTCAGTCCGGTGTCCACCAAAAGGTGGTCAATTTATGAA 2238
Qy 301 GCTGTGTAATGGGCTGAAATCTGCTACAAATCTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAATCTGCTACAAATCTTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
Qy 361 TGTCTACAGCAGCGCCAGACGGGAATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGAATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 2358
Qy 421 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCCGTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCCGTGAAGGTGTC 2478
Qy 541 GTTGAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGTCTCACACCCCTTGAGA 600
Db 2479 GTTGAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGTCTCACACCCCTTGAGA 2538
Qy 601 GAGCCGGTCAATGCAACGGAGACCGCTCTAGTGGCTACTATTCTACCAATATAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGGAGACCGCTCTAGTGGCTACTATTCTACCAATATAGATAT 2598
Qy 661 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
Qy 721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGTGAATGAGACAATA 2718
Qy 781 TATACAGTGGGAAGAGGCAATACACGGGAACCTTAATTTGGAGGTCAACCCCGAA 840
Db 2719 TATACAGTGGGAAGAGGCAATACACGGGAACCTTAATTTGGAGGTCAACCCCGAA 2778
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAAACCTCACTAGAAAA 2838
Qy 901 ATTGCAAGTGAAGAGTTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAAAATCATAGTGT 960
Db 2839 ATTGCAAGTGAAGAGTTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAAAATCATAGTGT 2898
Qy 961 CAGAGTCCGGCGGAATTTCTCCGACCCAGGACCAACACACACTGAAGACACACAA 1020
Db 2899 CAGAGTCCGGCGGAATTTCTCCGACCCAGGACCAACACACACTGAAGACACACAA 2958
Qy 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 3018
Qy 1081 GCAGTGTCCGATCTAACAACCCCTTGCCACAATCTCCACGAGTCCCCCAATCCCTCAACCC 1140
Db 3019 GCAGTGTCCGATCTAACAACCCCTTGCCACAATCTCCACGAGTCCCCCAATCCCTCAACCC 3078
Qy 1141 AAACCAAGTCCGGACACAGGACCCCAATATACACCGCTGTATAAATCTGCATCTCTGAG 1200
Db 3079 AAACCAAGTCCGGACACAGGACCCCAATATACACCGCTGTATAAATCTGCATCTCTGAG 3138

RESULT 14
ACC71549

ID ACC71549 standard; DNA; 8199 BP.

XX ACC71549;

XX 10-JUL-2003 (first entry)

XX VRC6604 (pAdApt Ebola GP(Z) (dTM)) plasmid.

XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;

KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;

KW viral infection; filovirus; circular; cyclic; ds.

OS Synthetic.

XX WO2003028632-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030251.

XX 01-OCT-2001; 2001US-0326476P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;

XX WPI; 2003-371961/35.

XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.

XX Claim 1; Page 182-185; 219pp; English.

XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected

CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence is an adenoviral shuttle vector
CC expressing Ebola virus glycoprotein (subtype Zaire) without its
CC transmembrane and intracellular domains
XX
SQ Sequence 8199 BP; 1977 A; 2047 C; 2173 G; 2002 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 10; Length 8199;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGCGCTTACAGGATATTGCGAGTACCTCGTGATCGAATTCAGAGGACATCATCTTT	60
Db	1336	ATGGCGCTTACAGGATATTGCGAGTACCTCGTGATCGAATTCAGAGGACATCATCTTT	1395
Qy	61	CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCATCCACCTTGGAGTCAACCAAT	120
Db	1396	CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCATCCACCTTGGAGTCAACCAAT	1455
Qy	121	AGCATTACAGGTTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA	180
Db	1456	AGCATTACAGGTTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA	1515
Qy	181	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	240
Db	1516	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	1575
Qy	241	TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGTCACCAAACTGAAATCAAAACCTGAGGAGTGAG	360
Db	1576	TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGTCACCAAACTGAAATCAAAATGAA	300
Qy	301	CTGTGTGAATGGCTGAAAACCTGTCTACAACTTGAATCAAAACCTGAGGAGTGAG	360
Db	1636	GCTGTGAATGGCTGAAAACCTGTCTACAACTTGAATCAAAACCTGAGGAGTGAG	1695
Qy	361	TGTTACAGAGGCCAGACGGGATTCGGGGCTTCCCGGTGTCGGGTATGTGACAAA	420
Db	1696	TGTTACAGAGGCCAGACGGGATTCGGGGCTTCCCGGTGTCGGGTATGTGACAAA	1755
Qy	421	GTATCAGGAACGGGACCGTGTGCGGAGACTTTCCTTCCATTAAGAGGCTTCTTCTTC	480
Db	1756	GTATCAGGAACGGGACCGTGTGCGGAGACTTTCCTTCCATTAAGAGGCTTCTTCTTC	1815
Qy	481	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC	540
Db	1816	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC	1875
Qy	541	GTTGCAATTTCTGATATCTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	600
Db	1876	GTTGCAATTTCTGATATCTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	1935
Qy	601	GAGCCGGTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCAATAGATAT	660
Db	1936	GAGCCGGTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCAATAGATAT	1995
Qy	661	CAGGCTACCGGTTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGTTGACAATTTGACC	720
Db	1996	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTGACAATTTGACC	2055
Qy	721	TACGTCCTCAATTTGAATCAAGATTTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA	780
Db	2056	TACGTCCTCAATTTGAATCAAGATTTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA	2115
Qy	781	TATACAGTGGGAAAAGAGGACATACCGAGGAAAACCTAATTTGGAAGGTCAACCCCGAA	840
Db	2116	TATACAGTGGGAAAAGAGGACATACCGAGGAAAACCTAATTTGGAAGGTCAACCCCGAA	2175
Qy	841	ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA	900
Db	2176	ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA	2235

Qy	901	ATTGCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACCGAGCCAAAAACATCAGTGGT	960
Db	2236	ATTGCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACCGAGCCAAAAACATCAGTGGT	2295
Qy	961	CAGAGTCCGGCGGCAACTTCTTCGAGCCAGGAGCAACAACTGTAAGACCAACAA	1020
Db	2296	CAGAGTCCGGCGGCAACTTCTTCGAGCCAGGAGCAACAACTGTAAGACCAACAA	2355
Qy	1021	ATCATGCTTCAGAAAAATTCCTTCGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1080
Db	2356	ATCATGCTTCAGAAAAATTCCTTCGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	2415
Qy	1081	GCAGTGTGCAATCTAAACAACCTTGGCCACATCTCCAGAGTCCCAATCCCTCACAACC	1140
Db	2416	GCAGTGTGCAATCTAAACAACCTTGGCCACATCTCCAGAGTCCCAATCCCTCACAACC	2475
Qy	1141	AAACAGTCCGGACAAACAGCACCACCAATAATACACCGTGTATAAATCTTGACATCTCTGAG	1200
Db	2476	AAACAGTCCGGACAAACAGCACCACCAATAATACACCGTGTATAAATCTTGACATCTCTGAG	2535
Qy	1201	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAAACAGACAGCAGCTCCGACACT	1260
Db	2536	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAAACAGACAGCAGCTCCGACACT	2595
Qy	1261	CCCTCTGCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACAGCAGCAAGAGC	1320
Db	2596	CCCTCTGCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACAGCAGCAAGAGC	2655
Qy	1321	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCAGCAGCCTGGC	1380
Db	2656	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCAGCAGCCTGGC	2715
Qy	1381	AACAACAACTCATCTCAAGATACCGGAGAAAGAGTGCCAGCAGCGGGAAGCTAGGC	1440
Db	2716	AACAACAACTCATCTCAAGATACCGGAGAAAGAGTGCCAGCAGCGGGAAGCTAGGC	2775
Qy	1441	TTAATTAACCAATCTATTTGCTGGAGTGCAGGACTGATCAAGCGGCGGAGAAAGACTCGA	1500
Db	2776	TTAATTAACCAATCTATTTGCTGGAGTGCAGGACTGATCAAGCGGCGGAGAAAGACTCGA	2835
Qy	1501	CGA 1503	
Db	2836	AGA 2838	
RESULT 15			
ID	ACC71548	standard; DNA; 8439 BP.	
XX	ACC71548;		
AC	ACC71548;		
DT	10-JUL-2003	(first entry)	
XX	VRC6603	(pAdapt Ebola GP(Z)) plasmid.	
XX	Virucide; vaccine; immune response; Ebola virus; Marburg virus;		
KW	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;		
KW	viral infection; filovirus; circular; cyclic; ds.		
OS	Synthetic.		
XX	WO2003028632-A2.		
PN	10-APR-2003.		
PD	24-SEP-2002; 2002WO-US030251.		
PF	01-OCT-2001; 2001US-0326476P.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA	Nabel GJ, Yang Z, Sullivan N, Sanchez A;		
PI			
XX			

DR WPI; 2003-371961/35.
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX Claim 1; Page 180-182; 219pp; English.
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX Sequence 8439 BP; 2059 A; 2081 C; 2210 G; 2089 T; 0 U; 0 Other;
SQ
Query Match 66.7%; Score 1501.4; DB 10; Length 8439;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGCTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATTTT 60
DB 1336 ATGGGGCTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATTTT 1395
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTATCCACAAAT 120
DB 1396 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTATCCACAAAT 1455
QY 121 ASCACATTACAGGTAGTATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
DB 1456 AGCACATTACAGGTAGTATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 1515
QY 181 AATCAATTGAGATCAGTTGGATCGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
DB 1516 AATCAATTGAGATCAGTTGGATCGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 1575
QY 241 TCTGCACTAAAGATGGGGCTTCAGGTCCGGTGTCCGACCAAGGTGTCAATATGAA 300
DB 1576 TCTGCACTAAAGATGGGGCTTCAGGTCCGGTGTCCGACCAAGGTGTCAATATGAA 1635
QY 301 GCTGTGTAATGGCTGAAACTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 360
DB 1636 GCTGTGTAATGGCTGAAACTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 1695
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGGCGGTATGTGCACAAA 420
DB 1696 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGGCGGTATGTGCACAAA 1755
QY 421 GTATCAGGAACGGGACCGGTGTCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480
DB 1756 GTATCAGGAACGGGACCGGTGTCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 1815
QY 481 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTCAAGGTGTC 540
DB 1816 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTCAAGGTGTC 1875
QY 541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA 600
DB 1876 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA 1935
QY 601 GAGCCGGTCAATGCAACGGAGGACCGGTCTAGTGGCTACTATTCTACCACAATTTAGATAT 660
DB 1936 GAGCCGGTCAATGCAACGGAGGACCGGTCTAGTGGCTACTATTCTACCACAATTTAGATAT 1995

QY 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTTGACAAATTTGACC 720
DB 1996 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTTGACAAATTTGACC 2055
QY 721 TACGTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 2056 TACGTCCAACTTGAATCAAGATTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 2115
QY 781 TATACAGTGGGAAAGGAGCAATACACCGGAAACCTAATTTGGAAGGTCAACCCCGAA 840
DB 2116 TATACAGTGGGAAAGGAGCAATACACCGGAAACCTAATTTGGAAGGTCAACCCCGAA 2175
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGCGAACTTAAAGAAACCTCCTCAGTAAAA 900
DB 2176 ATTGATACAACTCGGGAGTGGGCTTCTGCGAACTTAAAGAAACCTCCTCAGTAAAA 2235
QY 901 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTATCAACGGGAGCCAAACATCAGTGGT 960
DB 2236 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTATCAACGGGAGCCAAACATCAGTGGT 2295
QY 961 CAGAGTCCGCGCGCAACTTCTTCCGACCCAGGACCAACACAACTGAGACCCACAAA 1020
DB 2296 CAGAGTCCGCGCGCAACTTCTTCCGACCCAGGACCAACACAACTGAGACCCACAAA 2355
QY 1021 ATCATGGCTTCAGAAATTTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
DB 2356 ATCATGGCTTCAGAAATTTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 2415
QY 1081 GCAGTGTCCGATCTAAACAACTTCCGACCAATCTCCAGAGTCCCAATTCCTCACAACC 1140
DB 2416 GCAGTGTCCGATCTAAACAACTTCCGACCAATCTCCAGAGTCCCAATTCCTCACAACC 2475
QY 1141 AAACAGGTCCGCGACACAGACCCCAATTAATACCCGTGTATTAACCTTGACATCTCTGAG 1200
DB 2476 AAACAGGTCCGCGACACAGACCCCAATTAATACCCGTGTATTAACCTTGACATCTCTGAG 2535
QY 1201 GCACCTCAAGTTGAACAACTATCCGCAAGAACAGAACACAGACAGACCTCCGACACT 1260
DB 2536 GCACCTCAAGTTGAACAACTATCCGCAAGAACAGAACACAGACAGACCTCCGACACT 2595
QY 1261 CCCTTGTCCAGCAGCGCAGCGGACCCCAAAAGCAGAGAAACACAAACAGGAGCAGC 1320
DB 2596 CCCTTGTCCAGCAGCGCAGCGGACCCCAAAAGCAGAGAAACACAAACAGGAGCAGC 2655
QY 1321 ACTGACTTCTGGACCCCGCCACACACAAAGTCCCAAAACCCAGCAGAGACCCCTGGC 1380
DB 2656 ACTGACTTCTGGACCCCGCCACACAAAGTCCCAAAACCCAGCAGAGACCCCTGGC 2715
QY 1381 AACACAACTCATCAACAAAGATACCGGAGAAAGAGTGGCCAGCAGCGGGAAGCTAGGC 1440
DB 2716 AACACAACTCATCAACAAAGATACCGGAGAAAGAGTGGCCAGCAGCGGGAAGCTAGGC 2775
QY 1441 TTAATTACCAATATCTATTGTGGAGTCCAGAGCTGATCAGAGCGGGAGAAAGAACTCGA 1500
DB 2776 TTAATTACCAATATCTATTGTGGAGTCCAGAGCTGATCAGAGCGGGAGAAAGAACTCGA 2835
QY 1501 CGA 1503
DB 2836 AGA 2838

Search completed: July 31, 2005, 16:57:12
Job time : 1185.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 7351.67 seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-1
Perfect score: 2252
Sequence: 1 atgggcgtacaggaatt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	77	3.4	895	9	CNS0071A	AL066286 Drosophil
C 2	70.4	3.1	922	9	CNS0073W	AL066784 Drosophil
C 3	67.8	3.0	884	9	CNS006U0	AL065923 Drosophil
C 4	66.6	3.0	909	9	CNS00JTL	AL076720 Drosophil
C 5	62.6	2.8	1101	9	CNS00FXE	AL071170 Drosophil
C 6	60.4	2.7	902	9	CNS00EQP	AL065804 Drosophil
C 7	58.4	2.6	1101	9	CNS01J2T	AL080707 Drosophil
C 8	58.4	2.6	1225	9	CNS0166K	AL106358 Drosophil
C 9	57.2	2.5	918	9	AG405745	AG405745 Mus muscu
C 10	55.4	2.5	937	9	CNS006ST	AL065880 Drosophil
C 11	55.2	2.5	962	9	AG126141	AG126141 Pan trogl
C 12	54.4	2.4	939	9	CNS00CNG	AL059400 Drosophil
C 13	53.4	2.4	513	9	CNS02CMF	AL191328 Tetraodon
C 14	53.2	2.4	1806	9	CG756794	CG756794 P051-4-G0
C 15	53	2.4	934	9	CNS0004F	AL060213 Drosophil
C 16	53	2.4	1844	9	CG756580	CG756580 P051-4-B0
C 17	52.6	2.3	1051	9	CNS015H2	AL105440 Drosophil
C 18	52.6	2.3	1101	9	CNS001T2	AL078714 Drosophil
C 19	52.4	2.3	718	9	AG137045	AG137045 Pan trogl
C 20	52.2	2.3	788	9	AG125754	AG125754 Pan trogl
C 21	52.2	2.3	800	9	AG378374	AG378374 Mus muscu
C 22	52	2.3	733	9	CNS011AC	AL100014 Drosophil
C 23	51.6	2.3	626	6	BY752508	BY752508 BY752508
C 24	51.2	2.3	527	9	CE480293	CE480293 tigr-98s-

25	50.8	2.3	1101	9	CNS00LO0	AL068607 Drosophil
C 26	50.8	2.3	1101	9	CNS017RJ	AL108409 Drosophil
C 27	50.6	2.2	969	6	CD325655	CD325655 AGENCOURT
C 28	50.6	2.2	1101	9	CNS00Z1D	AL097099 Drosophil
C 29	50.4	2.2	1256	5	BQ433505	BQ433505 AGENCOURT
C 30	50.2	2.2	868	5	BU797633	BU797633 SJF2PC03
C 31	49.6	2.2	993	9	AG136567	AG136567 Pan trogl
C 32	49.6	2.2	1101	9	CNS017JU	AL108132 Drosophil
C 33	49.6	2.2	1752	9	CG756831	CG756831 P051-4-H0
C 34	49.4	2.2	572	9	CG897362	CG897362 pastbac04
C 35	49.4	2.2	790	9	CNS0379X	AL231054 Tetraodon
C 36	49.2	2.2	1022	8	BZ561449	BZ561449 pacs2-164
C 37	49.2	2.2	1087	9	CNS0106Y	AL098596 Drosophil
C 38	49	2.2	645	9	AG136205	AG136205 Pan trogl
C 39	48.8	2.2	1101	9	CNS017WI	AL108588 Drosophil
C 40	48.6	2.2	804	4	B1948870	B1948870 HVSME1001
C 41	48.6	2.2	926	8	AQ782232	AQ782232 HS_3176_B
C 42	48.4	2.1	1103	9	AG135574	AG135574 Pan trogl
C 43	48.4	2.1	1155	9	AG030839	AG030839 Pan trogl
C 44	48.2	2.1	712	9	CNS04AL2	AL281999 Tetraodon
C 45	48	2.1	714	9	CL855158	CL855158 OR_CBA008

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL066286
VERSION AL066286.1 GI:4945153
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)

ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE
1 (bases 1 to 895)
Phyloidea; Drosophilidae; Drosophila.

AUTHORS
DIRECT SUBMISSION

JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1. .895
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/note="end : TET3"

ORIGIN

Query Match 3.4%; Score 77; DB 9; Length 895;
Best Local Similarity 22.4%; Pred. No. 1.2e-10;

```
Matches 95; Conservative 150; Mismatches 180; Indels 0; Gaps 0;
QY 983 CCGAGCCGGGACCAACAACTGAAGACACACAAATATCTGGCTTCAGAAAATCTCT 1042
Db 894 CACGCCCCMMAMCCMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 835
QY 1043 CTGCAATGGTTCAAGTGCAAGTGAAGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1102
Db 834 MCMMMMAMMMCCMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 775
QY 1103 TTGCCACAATCTCCAGGAGTCCCAATCTCCATCAACCAACAGGTCGGGACACAGCA 1162
Db 774 MMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 715
QY 1163 CCATAATACACCCCGTGTAACTTGAATCTCTGAGGCAACTCAAGTTGAACAACTC 1222
Db 714 CCAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 655
QY 1223 ACCGAGAACAGACAGACAGACAGCTCCGACATCTCTCTGCGACGCGGAGCGG 1282
Db 654 MMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 595
QY 1283 GACCCCAAAAGCAGAGAACACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1342
Db 594 CACMMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 535
QY 1343 CCACAAAGTCCCAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1402
Db 534 MMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 475
QY 1403 ATACC 1407
Db 474 CMACC 470
```

```
RESULT 2
CNS0073W/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL066784
AL066784.1 GI:4945247
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 922)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1. .922
/organism="Drosophila melanogaster"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_lib="RPCI-98"
/note="End : TET3"

ORIGIN
Query Match 3.1%; Score 70.4; DB 9; Length 922;
Best Local Similarity 24.5%; Pred. No. 1.1e-08;
Matches 81; Conservative 116; Mismatches 133; Indels 0; Gaps 0;
QY 1088 CGCATCTTAACAAACCTTCCCAATCTCCAGAGTCCCAATCTCCATCAACAAACACAG 1147
Db 834 CCMACMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 775
QY 1148 GTCCGGACACAGCAGCACCATAATACACCGCTGTATAAATCTTGACATCTCTGAGGCAACTC 1207
Db 774 MCCACMMACACMMAMCCMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 715
QY 1208 AAGTTGAACAAATCACCGCAGAAACAGACAGACAGCAGCTCCGACATCTCCCTCTG 1267
Db 714 AAMMAAMMMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 655
QY 1268 CACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAGAGAGAGAGAGAGAG 1327
Db 654 ACAMMCMCACMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 595
QY 1328 TCCTGAGCCCGCCACCAACAAAGTCCCAAAACACAGCAGAGAGAGAGAGAGAGAG 1387
Db 594 ASACACAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMMAMMM 535
QY 1388 ACATCTCATCACCAAGATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
Db 534 CCMAAAAAAARAAAAAAGAAAAA 505
```

```
RESULT 3
CNS006U0/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL065923
AL065923.1 GI:4944891
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 884)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1. .884
```

```
FEATURES
source
```


REFERENCE	Ephydroidea; Drosophilidae; Drosophila.
AUTHORS	1 (bases 1 to 1101)
TITLE	Genoscope.
JOURNAL	Direct Submission
COMMENT	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

```

FEATURES
source
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37N10"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : T7"

```

[illegible]

RESULT 8	CNS0166K	linear	GSS 26-JUL-1999
LOCUS	CNS0166K/c	1225 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCSSION	AL106358		
VERSION	AL106358.1	GI:5621512	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydridea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1225)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : secret@genoscope.cns.fr)		
COMMENT	- Web : web.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) -		

http://www.edgp.ebi.ac.uk -. This *Drosophila melanogaster* BAC library (Dros BAC) was made by Alain Billaut at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genavieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES	Location/Qualifiers
source	1..1225
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACN15C20"
	/clone_lib="DrosBAC"
	/plasmid="pBelOBAC11"
	/note="end : Spe"

[illegible]

RESULT 9	AG405745	918 bp	DNA	linear	GSS 03-JUN-2004
LOCUS	AG405745/c				
DEFINITION	Mus musculus molossinus DNA, clone:MSMg01-264B16.TJ, genomic survey sequence.				
ACCESSION	AG405745				
VERSION	AG405745.1	GI:48048431			
KEYWORDS	GSS.				
SOURCE	Mus musculus molossinus				
ORGANISM	Mus musculus molossinus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1				
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
TITLE	BAC end Sequences of Library MSMg01				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 918)				
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
TITLE	Direct Submission				

JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT Clones are derived from the mouse BAC library MSMG01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : TJ

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

FEATURES **source**

1..918

Location/Qualifiers

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMG01-264B16.TJ"

/sex="male"

/tissue type="mixture of kidney and spleen"

/clone_lib="MSMG01 Mouse Male BAC Library"

ORIGIN

Query Match 2.5%; Score 57.2; DB 9; Length 918;

Best Local Similarity 52.1%; Pred. No. 8.6e-05;

Matches 148; Conservative 0; Mismatches 135; Indels 1; Gaps 1;

QY 1124 CCCATCCCTCACAACCAACAGGTCCGGACAACAGCACCCTATATACCCGGTGATA 1183
|||||
DB 865 CCCAGACCAACGACCCGCGCGGAAACACCAACACCAACCCCGCCGCATTA 806
|||||

QY 1184 AACTTGACATCTCTAGGCAACTCAAGTTGAAACAACATCACCGCAGACAGACAACGACA 1243
|||||

DB 805 CACCGACCAACCCCAACACCCACAGAACACACCATGATCCACACACAGATCCACA 746
|||||

QY 1244 GCACAGCTCCGACATCCCTCTGGCAGCAGCGCGGACCCCAAGCAGAGAACA 1303
|||||

DB 745 ACGACGCCCCACACAACCAACAATAGCTCGACCCACAGACCAACCCACCGCT 686
|||||

QY 1304 CCACACAGCAGCAGACGACTGCTCTGGACCCCGCCACCAACAGTCCCGAAACC 1363
|||||

DB 685 CCAACCCCAAGCGGACCACTCACTACCCAGATACGCGCCACCTAACCAACCCCC-CACCC 627
|||||

QY 1364 ACAGCGAGACCGCTGGCAACAACACATCTCATCCACCAAGATACC 1407
|||||

DB 626 ACAGCCACACCACTCCCCACACACCAACCAACCAACCAACCAAC 583
|||||

RESULT 10

CNS006ST 937 bp DNA linear GSS 03-JUN-1999

LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC14F16 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

DEFINITION AL065880

ACCESSION AL065880.1 GI:4944848

VERSION GSS

KEYWORDS Drosophila melanogaster (fruit fly)

SOURCE Drosophila melanogaster

ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscormorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 937)

AUTHORS Genoscope.

TITLE Direct Submision

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library](http://www.fruitfly.org/TheBDGP/Drosophila%20melanogaster%20BAC%20library) was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES **source**

1..937

Location/Qualifiers

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR14F16"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match 2.5%; Score 55.4; DB 9; Length 937;

Best Local Similarity 21.6%; Pred. No. 0.00029;

Matches 88; Conservative 140; Mismatches 179; Indels 0; Gaps 0;

QY 1092 TCTAACAAACCCCTTGCCACAATCTCCACGAGTCCCCAATCTCCACAACCAACACAGGTCC 1151
|||||

DB 437 TTTTYYMCMACMAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 496
|||||

QY 1152 GGACACAGCAGCACCATAATACACCGGTGTATAAATTGACATCTCTGAGGCACTCAAGT 1211
|||||

DB 497 MCMMACACCMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 556
|||||

QY 1212 TGACAACATCACCGCAAGACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1271
|||||

DB 557 CMMAAMCMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMA 616
|||||

QY 1272 GACCGCAGCCGAGCCCCCAAGCAGAGAACAACCAACAGCAGCAGCAGCAGCAGCAGCAGC 1331
|||||

DB 617 AAAAAAAMAAAMCM 676
|||||

QY 1332 GGACCCCGCCACCAACAAGTCCCCAAACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGC 1391
|||||

DB 677 MMCAAAVMVMGCAVCVMAARMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVM 736
|||||

QY 1392 TCATCACCAGATACCGGAGAGAGAGTGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1451
|||||

DB 737 AMAMCMRAAASAGASASRRRRGAACVVRGVGSARGSRGVVGGMRSGAGMARCAVMG 796
|||||

QY 1452 TACTATTGCTGGAGTCGAGGACTGATCAGACGGCGGAGAGAACTC 1498
|||||

DB 797 VAVRVASRAMRAASAGARVVMGVMGSAACRCVVMMAAAMAAHMTATM 843
|||||

RESULT 11

AGI26141

LOCUS AGI26141

DEFINITION Pan troglodytes DNA, clone: PTB-136J19.F, genomic survey sequence.

ACCESSION AGI26141

VERSION AGI26141.1 GI:16655306

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

JOURNAL Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,


```

10835645
2
REFERENCE
AUTHORS
  Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
  Fizames,C., Flecher,C., Bouneau,L., Billault,A., Quetier,F.,
  Saurin,W., Bernot,A. and Weissenbach,J.
TITLE
  Characterization and repeat analysis of the compact genome of the
  freshwater pufferfish Tetraodon nigroviridis
JOURNAL
  Genome Res. 10 (7), 939-949 (2000)
MEDLINE
  20359837
PUBMED
  10899143
REFERENCE
  3 (bases 1 to 513)
  Genoscope.
  Direct Submission
  Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
COMMENT
  This sequence is a single read and was generated as part of a large
  scale clone-end sequencing project of the Tetraodon nigroviridis
  genome. For more information, please take a look at
  http://www.genoscope.cns.fr/Tetraodon.
FEATURES
  Location/Qualifiers
    source
      1..513
      /organism="Tetraodon nigroviridis"
      /mol_type="genomic DNA"
      /db_xref="taxon:99883"
      /clone="255P15"
      /clone_lib="G"
      /note="Genoscope sequence ID : C0AG255CH08LPI-end : T7"
ORIGIN
  Query Match      2.4%; Score 53.4; DB 9; Length 513;
  Best Local Similarity 27.0%; Pred. No. 0.00094;
  Matches 84; Conservative 88; Mismatches 139; Indels 0; Gaps 0;
  QY 1100 CCCTTGCACATCTCCACAGTCCCAATCCCTCACAACCAACAGCGTCCGGACAACA 1159
  Db 444 CCMTTACTTTTTTTTTTAAACTTCCAAACANTCMAMCMCCAMAMWATWACAMCMWAMA 385
  QY 1160 GCACCATATACACCCGTGTATAACTTGACATCTCTGAGGCAACTCAAGTTGAACAAC 1219
  Db 384 AANTMMCMWAGAMCMCAACAAAAMAMCMCMACCMWATWAAAMMMWAGACMACWAS 325
  QY 1220 ATCACCGCAGACAGACGACGACGACGACGCTCCGACACTCCCTCTGCCAGCGCGAG 1279
  Db 324 ACCMCAWAGAMCMCMWATACCMWATMACAMAMWAGAMCMCMACMACCMWATWAA 265
  QY 1280 CCGGACCCCAAGAGCAGACACCAACACAGCAGACGAGACTGACTCTCTGGACCCCG 1339
  Db 264 AMACMCMTWAGAMCMCMACMACAMAMWAGAMCMCMWATWAAACAMWATWAMCMWMMG 205
  QY 1340 CCACCAACAAGTCCCAAAACCAACGAGCGAGCGCTGGCAACACACACTCATCACC 1399
  Db 204 AMAMMMWAGACCMATMACMCMCMACMASAMMMMMWAMCMCMCMYCASACCMWTCAS 145
  QY 1400 AAGATACCGGA 1410
  Db 144 ACACCACCASA 134
RESULT 14
CG756794/c
LOCUS
  CG756794 1806 bp DNA linear GSS 24-OCT-2003
DEFINITION
  P051-4-604.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
  genomic survey sequence.
ACCESSION
  CG756794
VERSION
  CG756794.1 GI:37984708
KEYWORDS
  GSS.
ORGANISM
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nemata; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
REFERENCE
  1 (bases 1 to 1806)
  Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
  Buntjer,J., van der Meulen,M. and Sommer,R.J.
  An integrated physical and genetic map of the nematode Pristionchus
  pacificus
  Mol. Genet. Genomics 269 (5), 715-722 (2003)
  22835951
  12884007
  COMMENT
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel: 00497071601371
  Fax: 00497071601498
  Email: ralf.sommer@tuebingen.mpg.de
  Class: BAC ends.
FEATURES
  Location/Qualifiers
    source
      1..1806
      /organism="Pristionchus pacificus"
      /mol_type="genomic DNA"
      /strain="California"
      /db_xref="taxon:54126"
      /clone_lib="Ppa EcoRI BAC Library"
      /note="The library was generated by a partial digest of
      the genomic DNA with EcoRI and cloning into the BAC
      vector."
ORIGIN
  Query Match      2.4%; Score 53.2; DB 9; Length 1806;
  Best Local Similarity 53.3%; Pred. No. 0.0016;
  Matches 153; Conservative 0; Mismatches 131; Indels 3; Gaps 2;
  QY 1113 CTCACGAGTCCCAATCCCTCACAACCAACAGAGTCCGGACACAGACCCATAATAC 1172
  Db 1610 CACCACAACACCGAACCACACAANCCACCACCAACAAACAAACACACACACCA 1551
  QY 1173 ACCCGTGTATAACTTGACATCTCTGAGGCAACTCAAGTTGAACAACATCACCGCAGAC 1232
  Db 1550 ACACCCACAACAACAACAAGACAGACACACCCACACACACNACAGCAGCACAACAAC 1491
  QY 1233 AGACAACGACAGCAGCGCTCCGACACTCTCTGCGACAGCGCGAGCGGACCCCAAA 1292
  Db 1490 AAACNAAACCAACACAC--ACCACACCCACCACCACACACACCGGCCCCCACC 1433
  QY 1293 AGCAGAGAACCAAC--CACGAGCAAGAGACTGACTTCTTGAGCCCGCAGCACAACNA 1351
  Db 1432 ACCCCACAACACCCACCACCAACACACCCGCGCACCACCCGCCACCCCAACAC 1373
  QY 1352 GTCCCCAAAACCAAGCAGCGCTGGCAACAACAACACTCATCAC 1398
  Db 1372 CCCCCCAACAGACCAACAACACCCCAAGCAGCACCACCCCAACACAC 1326
RESULT 15
CNS000D4F/c
LOCUS
  CNS000D4F 934 bp DNA linear GSS 04-JUN-1999
DEFINITION
  Drosophila melanogaster genome survey sequence TET3 end of BAC #
  BACR26H07 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL060213
VERSION
  AL060213.1 GI:4947565
KEYWORDS
  GSS.
SOURCE
  Drosophila melanogaster (fruit fly)
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 934)
  Genoscope.
  Direct Submission
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 360.705 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-1

Perfect score: 2252

Sequence: 1 atgggcgttacaggaattt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1501.4	66.7	2172	3	US-08-760-615-1
2	1499.8	66.6	2298	4	US-09-650-086A-1
3	742.2	33.0	11460	4	US-09-336-910A-1
4	740.2	32.9	2164	3	US-08-760-615-3
5	545.2	24.2	2247	3	US-08-760-615-5
6	49.2	2.2	399	4	US-09-621-976-8976
7	46.8	2.1	194889	4	US-09-949-016-15654
8	46.4	2.1	34230	4	US-09-949-016-12052
9	46.4	2.1	128470	4	US-09-949-016-13765
10	44.8	2.0	152132	4	US-09-949-016-13845
11	44.8	2.0	152145	4	US-09-949-016-12371
12	44.4	2.0	7218	1	US-08-232-463-14
13	44.2	2.0	30656	4	US-09-949-016-14613
14	43.8	1.9	522	4	US-09-949-016-103758
15	43.8	1.9	39154	4	US-09-949-016-12384
16	43.8	1.9	39154	4	US-09-949-016-12801
17	43.8	1.9	39443	4	US-09-949-016-14326
18	43.8	1.9	39443	4	US-09-949-016-14327
19	42	1.9	4403765	3	US-09-103-840A-2
20	42	1.9	4411529	3	US-09-103-840A-1
21	41.4	1.8	1968	4	US-09-248-796A-1729
22	41	1.8	242	4	US-09-354-147C-30
23	40.2	1.8	83617	4	US-09-949-016-12254
24	39.6	1.8	271134	4	US-09-949-016-12705
25	39.6	1.8	305491	4	US-09-949-016-17550
26	38.6	1.7	570	4	US-09-248-796A-7987
27	38.6	1.7	67620	4	US-09-949-016-16939

28	38.6	1.7	114793	4	US-10-148-806-3	Sequence 3, Appli
29	38.4	1.7	505	4	US-09-621-976-15639	Sequence 15639, A
30	38.4	1.7	1141	4	US-09-806-708B-22	Sequence 22, Appli
31	37.8	1.7	601	4	US-09-949-016-30710	Sequence 30710, A
32	37.8	1.7	601	4	US-09-949-016-182117	Sequence 182117,
33	37.8	1.7	3190	4	US-09-949-016-5219	Sequence 5219, Ap
34	37.8	1.7	3439	4	US-09-949-016-648	Sequence 648, App
35	37.8	1.7	31602	4	US-09-949-016-16961	Sequence 16961, A
36	37.4	1.7	2718	4	US-09-614-221A-376	Sequence 376, App
37	37.2	1.7	47981	4	US-09-679-279-1	Sequence 1, Appli
38	37	1.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
39	37	1.6	1059	4	US-09-248-796A-1415	Sequence 1415, Ap
40	37	1.6	1497	4	US-09-220-132-94	Sequence 94, Appl
41	36.8	1.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
42	36.8	1.6	13987	2	US-08-804-227C-13	Sequence 13, Appl
43	36.8	1.6	43280	2	US-08-804-227C-1	Sequence 1, Appli
44	36.6	1.6	486	4	US-09-639-207-13	Sequence 13, Appl
45	36.4	1.6	228851	4	US-09-949-016-13781	Sequence 13781, A

ALIGNMENTS

RESULT 1
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760.615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172 /product= "Glycoprotein"
; OTHER INFORMATION:

US-08-760-615-1

Query Match 66.7%; Score 1501.4; DB 3; Length 2172;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGTTACAGGAATATTGCAAGTATTGCGTACCTCGTGATCGAATCAAGAGGACATCATTTCTTT 60
DB |||||
QY 142 ATGGCGTTACAGGAATATTGCAAGTATTGCGTACCTCGTGATCGAATCAAGAGGACATCATTTCTTT 201
DB |||||
QY 61 CTTTGGTAAATATCTCTTTTCAAGAAACATTTTCCATCCCATCTTGGAGTCAATCACAAT 120
DB |||||
QY 202 CTTTGGTAAATATCTCTTTTCAAGAAACATTTTCCATCCCATCTTGGAGTCAATCACAAT 261
DB |||||
QY 121 AGCATTACAGGTAGTGTATGTCGCAAACTAGTTTGTGCGTACAACTGTCATCCACA 180
DB |||||
QY 262 AGCATTACAGGTAGTGTATGTCGCAAACTAGTTTGTGCGTACAACTGTCATCCACA 321
DB |||||
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
DB |||||
QY 322 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 381
DB |||||
QY 241 TCTGCACTAAAGATGGGGCTTCAGGTCGGTGTCCGACCAAGGTGTCAATATGAA 300
DB |||||
QY 382 TCTGCACTAAAGATGGGGCTTCAGGTCGGTGTCCGACCAAGGTGTCAATATGAA 441
DB |||||
QY 301 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAACCTGACGGGAGTGAG 360
DB |||||
QY 442 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAACCTGACGGGAGTGAG 501
DB |||||
QY 361 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAA 420
DB |||||
QY 502 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAA 561
DB |||||
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 480
DB |||||
QY 562 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 621
DB |||||
QY 481 CTGTATGATCGACTGCTTCCAGTTATCTACGAGGAACGACTTTCGCTCAAGGTGTC 540
DB |||||
QY 622 CTGTATGATCGACTGCTTCCAGTTATCTACGAGGAACGACTTTCGCTCAAGGTGTC 681
DB |||||
QY 541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA 600
DB |||||
QY 682 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAAGGACTTCTTCAGCTCACACCCCTTGAGA 741
DB |||||
QY 601 GAGCGGTCAATGCAACGGAGACCGCTGTAGTGCTACTATTCTACCAATATAGATAT 660
DB |||||
QY 742 GAGCGGTCAATGCAACGGAGACCGCTGTAGTGCTACTATTCTACCAATATAGATAT 801
DB |||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
DB |||||
QY 802 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 861
DB |||||
QY 721 TAGCTCCAACTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGAATGAGCAATA 780
DB |||||
QY 862 TAGCTCCAACTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGAATGAGCAATA 921
DB |||||
QY 781 TATCAAGTGGGAAAGGAGCAATACACGGGAAACTAATTTGGAGGTCAACCCCGAA 840
DB |||||
QY 922 TATCAAGTGGGAAAGGAGCAATACACGGGAAACTAATTTGGAGGTCAACCCCGAA 981
DB |||||
QY 841 ATTGATACAACTACGGGAGTGGGCTTCTGGGAACTTAAAGAACTCTACTAGAAA 900
DB |||||
QY 982 ATTGATACAACTACGGGAGTGGGCTTCTGGGAACTTAAAGAACTCTACTAGAAA 1041
DB |||||
QY 901 ATTTCAGGTGAAGAGTTGTCTTTTACAGTTGTATCAAGCGGACCCAAAACATCAGTGT 960
DB |||||
QY 1042 ATTTCAGGTGAAGAGTTGTCTTTTACAGTTGTATCAAGCGGACCCAAAACATCAGTGT 1101
DB |||||
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGACCAACAACTGAGGACCAAA 1020
DB |||||
QY 1102 CAGAGTCCGGCGGAACTTCTTCGACCCAGGACCAACAACTGAGGACCAAA 1161
DB |||||

QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB |||||
QY 1162 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1221
DB |||||
QY 1081 GCAGTGTGCAATCTTAAACAACTTGGCCCAATCTCCAGAGTCCCAATCCCTCAAC 1140
DB |||||
QY 1222 GCAGTGTGCAATCTTAAACAACTTGGCCCAATCTCCAGAGTCCCAATCCCTCAAC 1281
DB |||||
QY 1141 AAACAGGTCCGGCAACAGCAACCCATATACACCCGTGTATAACTTGACATCTCTGAG 1200
DB |||||
QY 1282 AAACAGGTCCGGCAACAGCAACCCATATACACCCGTGTATAACTTGACATCTCTGAG 1341
DB |||||
QY 1201 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGCAACAGCAGCAGCCTCCGACAT 1260
DB |||||
QY 1342 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGCAACAGCAGCAGCCTCCGACAT 1401
DB |||||
QY 1261 CCCTCTGCAGACCGCAGCGGACCCCAAGAGAGAACACCAACGAGGAGAGC 1320
DB |||||
QY 1402 CCCTCTGCAGACCGCAGCGGACCCCAAGAGAGAACACCAACGAGGAGAGC 1461
DB |||||
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCGCTGCG 1380
DB |||||
QY 1462 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCGCTGCG 1521
DB |||||
QY 1381 AACAAACAACTCATCACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGAGAGCTAGGC 1440
DB |||||
QY 1522 AACAAACAACTCATCACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGAGAGCTAGGC 1581
DB |||||
QY 1441 TTAATTAACAATATCTATTGCTGGAGTCCGAGACTGTATCAAGGCGGAGAGAACTCGA 1500
DB |||||
QY 1582 TTAATTAACAATATCTATTGCTGGAGTCCGAGACTGTATCAAGGCGGAGAGAACTCGA 1641
DB |||||
QY 1501 CGA 1503
DB 1642 AGA 1644

RESULT 2

US-09-650-086A-1
; Sequence 1, Application US/09650086A
; Patent No. 6630144
; GENERAL INFORMATION:
; APPLICANT: Hart, Mary Kate
; APPLICANT: Wilson, Julie A.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Prophylactic and Therapeutic Monoclonal Antibodies
; FILE REFERENCE: 003/198/SAP
; CURRENT APPLICATION NUMBER: US/09/650,086A
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: US 60/151,505
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola Zaire
US-09-650-086A-1

Query Match 66.6%; Score 1499.8; DB 4; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGTTACAGGAATATTGCAAGTATTGCGTACCTCGTGATCGAATCAAGAGGACATCATTTCTTT 60
DB |||||
QY 59 ATGGCGTTACAGGAATATTGCAAGTATTGCGTACCTCGTGATCGAATCAAGAGGACATCATTTCTTT 118
DB |||||
QY 61 CTTTGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
DB |||||
QY 119 CTTTGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 178
DB |||||
QY 121 AGCATTACAGGTAGTGTATGTCGACAAACTAGTTTGTGTCGACAAACTGTCATCCACA 180
DB |||||

Db 179 AGCATTACAGTTAGTGTGACAACTAGTTTGTGTGACAACTGTCTATCCACA 238
Qy 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 239 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
Qy 241 TCTGCAATTAAGATGGGCTTCAGGTCGGGTCCGCTCCACCAAGGTTGGTCAATATGAA 300
Db 299 TCTGCAATTAAGATGGGCTTCAGGTCGGGTCCGCTCCACCAAGGTTGGTCAATATGAA 358
Qy 301 GCTGTGAATGGGCTGAAAACCTGTCTACAACTCTTGAAATCAAAAACCTGACGGAGTGAG 360
Db 359 GCTGTGAATGGGCTGAAAACCTGTCTACAACTCTTGAAATCAAAAACCTGACGGAGTGAG 418
Qy 361 TGTCTACAGAGCCGACAGACGGGATTCGGGCTTCGCCGTTCGCCGTATGTGCACAAA 420
Db 419 TGTCTACAGAGCCGACAGACGGGATTCGGGCTTCGCCGTTCGCCGTATGTGCACAAA 478
Qy 421 GTATCAGGAAACGGGACCGTGTGCGGAGACTTTCGCTTCCATTAAGAGGTTGCTTCTTC 480
Db 479 GTATCAGGAAACGGGACCGTGTGCGGAGACTTTCGCTTCCATTAAGAGGTTGCTTCTTC 538
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGTGTGC 540
Db 539 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGTGTGC 598
Qy 541 GTTGCAATTTCTGATCTGCCCAAGCTAAGAGGACTTCTTCACTCAGCTCACACCCCTTGAGA 600
Db 599 GTTGCAATTTCTGATCTGCCCAAGCTAAGAGGACTTCTTCACTCAGCTCACACCCCTTGAGA 658
Qy 601 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGTGCTACTATTCTACCAATATAGATAT 660
Db 659 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGTGCTACTATTCTACCAATATAGATAT 718
Qy 661 CAGGCTACCGGTTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 719 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 778
Qy 721 TACGTCCTCACTTGAATCAAGATTTCAACCAAGTCTTCTGCTCCAGCTGAATGAGACAATA 780
Db 779 TACGTCCTCACTTGAATCAAGATTTCAACCAAGTCTTCTGCTCCAGCTGAATGAGACAATA 838
Qy 781 TATACAGTGGGAAAGAGGCAATACACGGGAAACCTAATTTGGAAGTCAACCCCGAA 840
Db 839 TATACAGTGGGAAAGAGGCAATACACGGGAAACCTAATTTGGAAGTCAACCCCGAA 898
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACCTAATAAAACCTCACTAGAAA 900
Db 899 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACCTAATAAAACCTCACTAGAAA 958
Qy 901 ATTCGAGTGAAGATTTGTCTTTTACAGTTGTATCAAAACGGGACCAAAAACATCAGTGGT 960
Db 959 ATTCGAGTGAAGATTTGTCTTTTACAGTTGTATCAAAACGGGACCAAAAACATCAGTGGT 1018
Qy 961 CAGAGTCCGGCCGCAACTTTCTTCGACCGAGGACCAACAACTGAGGACCAACAAA 1020
Db 1019 CAGAGTCCGGCCGCAACTTTCTTCGACCGAGGACCAACAACTGAGGACCAACAAA 1078
Qy 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 1079 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1138
Qy 1081 GCAGTGTGCACTTCAACACCCCTTGCACAACTCTCCAGAGTCCCAATCTCCACAC 1140
Db 1139 GCAGTGTGCACTTCAACACCCCTTGCACAACTCTCCAGAGTCCCAATCTCCACAC 1198
Qy 1141 AAACAGGTCGGGACCAACAGCACCATAATACACCGGTGTATPAACTTGACATCTCTGAG 1200
Db 1199 AAACAGGTCGGGACCAACAGCACCATAATACACCGGTGTATPAACTTGACATCTCTGAG 1258
Qy 1201 GCAACTCAAGTTGAACACATCACCGGAAACAGACAGCAGCAGCAGCCTCCGACACT 1260

Db 1259 GCAACTCAAGTTGAACAAGATCATCCGCGAGAACACACACAGCAGCCTCCGACACT 1318
Qy 1261 CCCTCTGCCAGCAGCCGAGCGGACCCCAAAAGAGAGAAACCAACACAGAGCAAGAGC 1320
Db 1319 CCCTCTGCCAGCAGCCGAGCGGACCCCAAAAGAGAGAAACCAACACAGAGCAAGAGC 1378
Qy 1321 ACTGACTTCTGGACCCCGCCACCAACACAGTCCCAAAACCAACAGCAGCAGCAGCCTGCGC 1380
Db 1379 ACTGACTTCTGGACCCCGCCACCAACACAGTCCCAAAACCAACAGCAGCAGCAGCCTGCGC 1438
Qy 1381 AACCAACACTCATCATCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 1439 AACCAACACTCATCATCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1498
Qy 1441 TTAATTAACCAATATCTATTGCTGGAGTGCAGGACTGATCAAGCGGAGAGAACTCGA 1500
Db 1499 TTAATTAACCAATATCTATTGCTGGAGTGCAGGACTGATCAAGCGGAGAGAACTCGA 1558
Qy 1501 CGA 1503
Db 1559 AGA 1561

RESULT 3
US-09-336-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan P.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-09-336-910A-1

Query Match 33.0%; Score 742.2; DB 4; Length 11460;
Best Local Similarity 98.9%; Pred. No. 7.3e-226;
Matches 747; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1492 AGAATCGACCATCGATCCTCTGGAGGGAAGCGGACATGTTCCCTTTCTGGATGGTTA 1551
Db 7233 AGAAGAAAGCAAGATCCTCTGGAGGGAAGCGGACATGTTCCCTTTCTGGATGGTTA 7292
Qy 1552 ATAAATGCTCCAATTTGATTTTGACCCAGTTCCAAATACAAAACAACTCTTTGATGAATCC 1611
Db 7293 ATAAATGCTCCAATTTGATTTTGACCCAGTTCCAAATACAAAACAACTCTTTGATGAATCC 7352
Qy 1612 TCTAGTTCGTGCTCGGCTGAGGAAGATCAACATGCTCCCAATATTAGTTTAATC 1671
Db 7353 TCTAGTTCGTGCTCGGCTGAGGAAGATCAACATGCTCCCAATATTAGTTTAATC 7412
Qy 1672 TTATCTTATTTTCTTAATATAATGAGAACTGCTCTACTCTGGAGAAATGAGAAATGAT 1731
Db 7413 TTATCTTATTTTCTTAATATAATGAGAACTGCTCTACTCTGGAGAAATGAGAAATGAT 7472
Qy 1732 TGTGATGAGAGTAAAGAAATTTGGAGGCTTCAGAGGATGACCTGCGCGCAGGGCTCAGT 1791
Db 7473 TGTGATGAGAGTAAAGAAATTTGGAGGCTTCAGAGGATGACCTGCGCGCAGGGCTCAGT 7532

```
QY 1792 TGGATACCCGTTTTTGGCCCTCGAATTGAAGACATTTACATGCTGTTTAAATTTAAAAAT 1851
Db 7533 TGGATACCCGTTTTTGGCCCTCGAATTGAAGACATTTACATGCTGTTTAAATTTAAAAAT 7592
QY 1852 CAAAACAAATTTGGTCTGCAGGTTGAGGCGCTAGCCCAATCAAACTGCCAAATCCTTGGAA 1911
Db 7593 CAAAACAAATTTGGTCTGCAGGTTGAGGCGCTAGCCCAATCAAACTGCCAAATCCTTGGAA 7652
QY 1912 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 1971
Db 7653 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 7712
QY 1972 GACTTTCTACTCACAAGATGGGAGGAAACATCAAAAGTCTTTGGACCTGATTGTTGCATC 2031
Db 7713 GACTTTCTACTCACAAGATGGGAGGAAACATCAAAAGTCTTTGGACCTGATTGTTGCATC 7772
QY 2032 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATGACCAAAATTTAAAGAGACGAA 2091
Db 7773 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATGACCAAAATTTAAAGAGACGAA 7832
QY 2092 CAAAAGAGGGGACTGGTTGGGCTGCGGTGGTAAATGGTGACATCCGACTGGGGTGT 2151
Db 7833 CAAAAGAGGGGACTGGTTGGGCTGCGGTGGTAAATGGTGACATCCGACTGGGGTGT 7892
QY 2152 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2211
Db 7893 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 7952
QY 2212 ATTTGCTGATCTTTACTTAATATATATCGGATAACG 2246
Db 7953 ATTTGCTGATCTTTACTTAATATATATCGGATAACG 7987
```

RESULT 4

```
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Mueoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3
```

```
Query Match 32.9%; Score 740.2; DB 3; Length 2164;
Best Local Similarity 98.9%; Pred. No. 9.2e-226;
Matches 745; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1492 AGAATCCGACCATCGATCCTCTGGAGGGAAGCGACATGTTCCCTTTCTGATGGGTGA 1551
Db 1412 AGAAGAAAGCAAGATCCTCTGGAGGGAAGCGACATGTTCCCTTTCTGATGGGTGA 1471
QY 1552 ATAAATGCTCCAAATTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1611
Db 1472 ATAAATGCTCCAAATTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1531
QY 1612 TCTAGTTCTGTGCTCGGCTGAGGAAGATCAACATGCTCCCTCCCAATATTAGTTTAACT 1671
Db 1532 TCTAGTTCTGTGCTCGGCTGAGGAAGATCAACATGCTCCCTCCCAATATTAGTTTAACT 1591
QY 1672 TTATCTTAATTTTCTTAATATAATGAGAACACTGCTACTCTGGAGAAAAAGAGAAATGAT 1731
Db 1592 TTATCTTAATTTTCTTAATATAATGAGAACACTGCTACTCTGGAGAAAAAGAGAAATGAT 1651
QY 1732 TGTGATGCAGAGTTAAAGAAATTTGAGAGCGTTCAGGAGGATGACCTGCGCGCAGGGCTCAGT 1791
Db 1652 TGTGATGCAGAGTTAAAGAAATTTGAGAGCGTTCAGGAGGATGACCTGCGCGCAGGGCTCAGT 1711
QY 1792 TGGATACCCGTTTTTGGCCCTGGAATTTGAAGGACATTTACATGCTGTTTAAATTTAAAAAT 1851
Db 1712 TGGATACCCGTTTTTGGCCCTGGAATTTGAAGGACATTTACATGCTGTTTAAATTTAAAAAT 1771
QY 1852 CAAAACAAATTTGGTCTGCAGGTTGAGGCGCTAGCCCAATCAAACTGCCAAATCCTTGGAA 1911
Db 1772 CAAAACAAATTTGGTCTGCAGGTTGAGGCGCTAGCCCAATCAAACTGCCAAATCCTTGGAA 1831
QY 1912 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 1971
Db 1832 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 1891
QY 1972 GACTTTCTACTCACAAGATGGGAGGAAACATGCAAAAGTCTTTGGACCTGATTGTTGCATC 2031
Db 1892 GACTTTCTACTCACAAGATGGGAGGAAACATGCAAAAGTCTTTGGACCTGATTGTTGCATC 1951
QY 2032 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATTTGACCAAAATTTAAAGAGACGAA 2091
Db 1952 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATTTGACCAAAATTTAAAGAGACGAA 2011
QY 2092 CAAAAGAGGGGACTGGTTGGGCTGCGGTGTAATGTTGACATCCGACTGGGGTGT 2151
Db 2012 CAAAAGAGGGGACTGGTTGGGCTGCGGTGTAATGTTGACATCCGACTGGGGTGT 2071
QY 2152 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2211
Db 2072 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2131
QY 2212 ATTTGCTGATCTTTACTTAATATATATCGGATAA 2244
Db 2132 ATTTGCTGATCTTTACTTAATATATATCGGATAA 2164
```

RESULT 5

```
US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1415.31 Seconds
(without alignment)
10298.647 Million cell updates/sec

Title: US-10-066-506A-1
Perfect score: 2252
Sequence: 1 atgggcgttacaggaattt.....tatatcgataacggaatttc 2252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2252	100.0	2252	15 US-10-066-506A-1	Sequence 1, Appli
2	1506	66.9	2039	15 US-10-066-506A-11	Sequence 11, Appli
3	1501.4	66.7	2298	18 US-10-226-795-1	Sequence 1, Appli
4	1501.4	66.7	6987	20 US-10-491-121-6	Sequence 6, Appli
5	1501.4	66.7	6914	20 US-10-491-121-9	Sequence 9, Appli
6	1501.4	66.7	7044	20 US-10-491-121-7	Sequence 7, Appli
7	1501.4	66.7	7106	20 US-10-491-121-8	Sequence 8, Appli

8	1501.4	66.7	7154	20	US-10-491-121-1	Sequence 1, Appli
9	1501.4	66.7	7188	20	US-10-491-121-2	Sequence 2, Appli
10	1501.4	66.7	7285	22	US-10-860-878-3	Sequence 3, Appli
11	1501.4	66.7	8199	20	US-10-491-121-29	Sequence 28, Appli
12	1501.4	66.7	8439	20	US-10-491-121-28	Sequence 28, Appli
13	1501.4	66.7	10783	20	US-10-491-121-25	Sequence 25, Appli
14	1499.8	66.6	2298	9	US-09-337-946A-1	Sequence 1, Appli
15	1499.8	66.6	2298	17	US-10-384-976-1	Sequence 1, Appli
16	1499.8	66.6	2298	22	US-10-696-633-1	Sequence 1, Appli
17	1489.4	66.1	7272	22	US-10-860-878-4	Sequence 4, Appli
18	1489.4	66.1	18959	17	US-10-353-856-19	Sequence 19, Appli
19	1489.4	66.1	18959	17	US-10-353-856-37	Sequence 37, Appli
20	1489.4	66.1	18959	17	US-10-353-856-46	Sequence 46, Appli
21	1480	65.7	6724	20	US-10-491-121-5	Sequence 5, Appli
22	1446.2	64.2	2030	21	US-10-811-353-4	Sequence 4, Appli
23	906.4	40.2	6624	20	US-10-491-121-3	Sequence 3, Appli
24	905.4	40.2	6561	20	US-10-491-121-4	Sequence 4, Appli
25	900	40.0	6324	20	US-10-491-121-37	Sequence 37, Appli
26	753	33.4	2051	15	US-10-066-506A-9	Sequence 9, Appli
27	746.6	33.2	2046	15	US-10-066-506A-7	Sequence 7, Appli
28	742.2	33.0	7778	20	US-10-491-121-30	Sequence 30, Appli
29	742.2	33.0	11460	16	US-10-267-322-1	Sequence 1, Appli
30	685.8	30.5	2046	21	US-10-811-353-2	Sequence 2, Appli
31	685.8	30.5	19112	17	US-10-353-856-27	Sequence 27, Appli
32	675.2	30.0	6467	20	US-10-491-121-10	Sequence 10, Appli
33	644.2	28.6	7005	20	US-10-491-121-31	Sequence 31, Appli
34	644.2	28.6	8256	20	US-10-491-121-32	Sequence 32, Appli
35	548.2	24.3	2046	15	US-10-066-506A-13	Sequence 13, Appli
36	546.6	24.3	2046	15	US-10-066-506A-5	Sequence 5, Appli
37	546.2	24.3	6885	20	US-10-491-121-18	Sequence 18, Appli
38	546.2	24.3	7002	20	US-10-491-121-16	Sequence 16, Appli
39	546.2	24.3	7023	20	US-10-491-121-21	Sequence 21, Appli
40	546.2	24.3	7036	20	US-10-491-121-17	Sequence 17, Appli
41	534.2	23.7	6889	20	US-10-491-121-19	Sequence 19, Appli
42	534.2	23.7	7003	22	US-10-860-878-1	Sequence 1, Appli
43	534.2	23.7	8146	20	US-10-491-121-20	Sequence 20, Appli
44	466.2	20.7	18890	17	US-10-353-856-9	Sequence 9, Appli
45	464.6	20.6	2360	17	US-10-397-635-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-10-066-506A-1
; Sequence 1, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 2
; OTHER INFORMATION: and Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-1

Query Match 100.0%; Score 2252; DB 15; Length 2252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGTACAGGAATTGCTACCTCGTGATCGATTTCAGAGGACATCTTT 60

|||||
1 ATGGCGTTTACAGGAATATTGAGTTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 60
QY
61 CTTTGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
Db
61 CTTTGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
QY
121 AGCACATTACAGGTTAGTGTGATGTCGACAAATAGTTTGTGTCGACAAATCTGTCATCCACA 180
Db
121 AGCACATTACAGGTTAGTGTGATGTCGACAAATAGTTTGTGTCGACAAATCTGTCATCCACA 180
QY
181 AATCAATTGAGATCAGTTGGAATCTGGAATCTGGAAGGAATGGAAGTGGCAACTGTCATGCA 240
Db
181 AATCAATTGAGATCAGTTGGAATCTGGAATCTGGAAGGAATGGAAGTGGCAACTGTCATGCA 240
QY
241 TCTGCAACTTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAAAGGTGGTCAATTTATGNA 300
Db
241 TCTGCAACTTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAAAGGTGGTCAATTTATGNA 300
QY
301 GCTGTGAAATGGGCTGAAACTGCTACAAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db
301 GCTGTGAAATGGGCTGAAACTGCTACAAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
QY
361 TGTCTACCAAGCGGCAGACGGGATTCGGGGCTTCCCGGTGCGGTGATGTGCACAAA 420
Db
361 TGTCTACCAAGCGGCAGACGGGATTCGGGGCTTCCCGGTGCGGTGATGTGCACAAA 420
QY
421 GTATCAGGAACGGGACCGGTGCGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 480
Db
421 GTATCAGGAACGGGACCGGTGCGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 480
QY
481 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTCCGCTGAAAGTGTG 540
Db
481 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTCCGCTGAAAGTGTG 540
QY
541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCACTGTCACACCCCTTGAGA 600
Db
541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCACTGTCACACCCCTTGAGA 600
QY
601 GAGCCGGTCAATGCAACGAGGACCGCTAGTGGCTACTATCTACCAATTAGATAT 660
Db
601 GAGCCGGTCAATGCAACGAGGACCGCTAGTGGCTACTATCTACCAATTAGATAT 660
QY
661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTGAGGTTGACAAATTTGACC 720
Db
661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTGAGGTTGACAAATTTGACC 720
QY
721 TAGCTCCAACTTGAATCAAGATTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db
721 TAGCTCCAACTTGAATCAAGATTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
QY
781 TATACAAGTGGGAAAGGAGCAATACCGGGAAACATAATTTGGAGGTCAACCCCGAA 840
Db
781 TATACAAGTGGGAAAGGAGCAATACCGGGAAACATAATTTGGAGGTCAACCCCGAA 840
QY
841 ATTGATACAACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 900
Db
841 ATTGATACAACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 900
QY
901 ATTTCGAGTGAAGAGTTGTCTTTTCACAGTTGTATCAAAACGGGACCAAAACATCAAGTGT 960
Db
901 ATTTCGAGTGAAGAGTTGTCTTTTCACAGTTGTATCAAAACGGGACCAAAACATCAAGTGT 960
QY
961 CAGAGTCCGGCGGAACTTCTTCCGACCCAGGGACCAACACAACTCAAGACCAACAA 1020
Db
961 CAGAGTCCGGCGGAACTTCTTCCGACCCAGGGACCAACACAACTCAAGACCAACAA 1020
QY
1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGAGCT 1080
Db
1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGAGCT 1080
QY
1081 GCAGTGTGCAATCAACAAACCTTTGCGCAAAATCTTCCACGAGTCCCAATCCCTCAACCC 1140
|||||

Db
1081 GCAGTGTGCAATCAACAAACCTTTGCGCAAAATCTCCACGAGTCCCAATCCCTCAACACC 1140
QY
1141 AAACCCAGTCCGGCAACACAGCACCCATTAATACACCCGTTGATAAACTTTGACATCTCTGAG 1200
Db
1141 AAACCCAGTCCGGCAACACAGCACCCATTAATACACCCGTTGATAAACTTTGACATCTCTGAG 1200
QY
1201 GCAACTCAAGTTGAACAACTATCAGCGCAAGAACAGAACAGACAGACAGCTCCGACACT 1260
Db
1201 GCAACTCAAGTTGAACAACTATCAGCGCAAGAACAGAACAGACAGACAGCTCCGACACT 1260
QY
1261 CCCTCTGCCACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 1320
Db
1261 CCCTCTGCCACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 1320
QY
1321 ACTGACTTCTGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 1380
Db
1321 ACTGACTTCTGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 1380
QY
1381 AACAAACAACTCATCATCAACCAAGATACCGGAGAGAGAGTGCAGACGCGGAGAGTGGC 1440
Db
1381 AACAAACAACTCATCATCAACCAAGATACCGGAGAGAGAGTGCAGACGCGGAGAGTGGC 1440
QY
1441 TTAATTACCAATACTATTTGCTGGAGTCCGAGGACTGATCACAGCGGAGAGAGTGGC 1500
Db
1441 TTAATTACCAATACTATTTGCTGGAGTCCGAGGACTGATCACAGCGGAGAGAGTGGC 1500
QY
1501 CGATCGATCTCTGGAGGAGGAGGACATGTTCCCTTTCTGGATGGGTAAATAATGCT 1560
Db
1501 CGATCGATCTCTGGAGGAGGAGGAGGACATGTTCCCTTTCTGGATGGGTAAATAATGCT 1560
QY
1561 CCAATTGATTTTGAACCCAGTTTCCAAATAACAAACAAATCTTTTGATGAATCTCTAGTCT 1620
Db
1561 CCAATTGATTTTGAACCCAGTTTCCAAATAACAAACAAATCTTTTGATGAATCTCTAGTCT 1620
QY
1621 GGTGCTCGCTGAGGAGAGATCAATGCTCTCCCAATATAGTTTAACTTATCTTAT 1680
Db
1621 GGTGCTCGCTGAGGAGAGATCAATGCTCTCCCAATATAGTTTAACTTATCTTAT 1680
QY
1681 TTTCTTAATATAAATGAGAACACTGCTTACTCTGAGAGAAATGAGAAATGATGTGATGCA 1740
Db
1681 TTTCTTAATATAAATGAGAACACTGCTTACTCTGAGAGAAATGAGAAATGATGTGATGCA 1740
QY
1741 GAGTTAAGAAATTTGAGGCGTTTCAAGAGGATGACCTGCGCGGAGGCTCAGTTGGATACCG 1800
Db
1741 GAGTTAAGAAATTTGAGGCGTTTCAAGAGGATGACCTGCGCGGAGGCTCAGTTGGATACCG 1800
QY
1801 TTTTGTGCCCCGAAATTTGAAGGACTTTTACACTGCTGTTTAAATTTAAATCAAAACAA 1860
Db
1801 TTTTGTGCCCCGAAATTTGAAGGACTTTTACACTGCTGTTTAAATTTAAATCAAAACAA 1860
QY
1861 TTGGTCTCAGGTTGAGGCGTCTAGCCAAATCAAACTGCAAAATCTTGGAACTCTTATTTG 1920
Db
1861 TTGGTCTCAGGTTGAGGCGTCTAGCCAAATCAAACTGCAAAATCTTGGAACTCTTATTTG 1920
QY
1921 AGAGTCAAACTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTTCTA 1980
Db
1921 AGAGTCAAACTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTTCTA 1980
QY
1981 CTCACAAATGAGGAGGAGAACATGCAAGTGTGAGCCTGATTTGATCGGATAGAA 2040
Db
1981 CTCACAAATGAGGAGGAGAACATGCAAGTGTGAGCCTGATTTGATCGGATAGAA 2040
QY
2041 GACTTGTCCAAAAATTTTTCAGAGCAAAATTTGACCAAAATTTAAAAAGGAGCAAAAGAG 2100
Db
2041 GACTTGTCCAAAAATTTTTCAGAGCAAAATTTGACCAAAATTTAAAAAGGAGCAAAAGAG 2100
QY
2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTCTTCTACTAAC 2160
Db
2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTCTTCTACTAAC 2160
QY
2161 TTGGGCATTTTGTCTACTATTTATCCATAGTGTCTGATTTGATCTCTGATTTTGTCTGT 2220
Db
2161 TTGGGCATTTTGTCTACTATTTATCCATAGTGTCTGATTTGATCTCTGATTTTGTCTGT 2220

QY 2221 ATCTTTACTAATATATATCGGATAACGGAATTC 2252
|||||
Db 2221 ATCTTTACTAATATATATCGGATAACGGAATTC 2252
|||||

RESULT 2

US-10-066-506a-11
; Sequence 11, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 11
; LENGTH: 2039
; TYPE: DNA
; ORGANISM: Ebola virus Zaire strain
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Glycoprotein 1 and Ebola vi
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506a-11

Query Match 66.9%; Score 1506; DB 15; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATGCGATTGCTGATCGATTCAAGAGGACATCTTTT 60
|||||
Db 1 ATGGCGGTTACAGGAATATGCGATTGCTGATCGATTCAAGAGGACATCTTTT 60
|||||
QY 61 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT 120
|||||
Db 61 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT 120
|||||
QY 121 AGCATTACAGGTTAGTGTGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
|||||
Db 121 AGCATTACAGGTTAGTGTGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
|||||
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 240
|||||
Db 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 240
|||||
QY 241 TCTGCAACTAAAGATGGGCTTCAAGGTCGGGTGTCACCAAAAGGTGGTCAATTTATGAA 300
|||||
Db 241 TCTGCAACTAAAGATGGGCTTCAAGGTCGGGTGTCACCAAAAGGTGGTCAATTTATGAA 300
|||||
QY 301 GCTGTGTAATGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTAG 360
|||||
Db 301 GCTGTGTAATGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTAG 360
|||||
QY 361 TGCTTACAGCAGCGCCAGAGCGGATTCGGGGCTTCCCGCGGTGCGGTATGTCACAAA 420
|||||
Db 361 TGCTTACAGCAGCGCCAGAGCGGATTCGGGGCTTCCCGCGGTGCGGTATGTCACAAA 420
|||||
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAGAGGGTCTTTCTTC 480
|||||
Db 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAGAGGGTCTTTCTTC 480
|||||
QY 481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
|||||
Db 481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
|||||
QY 541 GTTGCATTTCTGATCTGCCCCCAAGCTAAGAGGACTTCTCAGCTCACACCCCTTGAGA 600
|||||

RESULT 3
US-10-226-795-1
; Sequence 1, Application US/10226795
; Publication No. US20040053865A1
; GENERAL INFORMATION:
; APPLICANT: HART, MARY KATE

Db 541 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTTCAGCTCACACCCCTTGAGA 600
QY 601 GAGCCGTCATATGCAACGAGAGCCCGCTAGTGGCTACTATTTCTACCAATTTAGATAT 660
|||||
Db 601 GAGCCGTCATATGCAACGAGAGCCCGCTAGTGGCTACTATTTCTACCAATTTAGATAT 660
|||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGGTTGCAATTTGACC 720
|||||
Db 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGGTTGCAATTTGACC 720
|||||
QY 721 TACGTCCAACTTGAATCAAGATTCACACCAAGTTTCTGTCTCCAGCTGAAATGAGACAATA 780
|||||
Db 721 TACGTCCAACTTGAATCAAGATTCACACCAAGTTTCTGTCTCCAGCTGAAATGAGACAATA 780
|||||
QY 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAATAATTTTGGAGGTCAACCCGAA 840
|||||
Db 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAATAATTTTGGAGGTCAACCCGAA 840
|||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAAATAATTTTGGAGGTCAACCCGAA 900
|||||
Db 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAAATAATTTTGGAGGTCAACCCGAA 900
|||||
QY 901 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACGAGGCCAAAAACATCAGTGT 960
|||||
Db 901 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAAAACGAGGCCAAAAACATCAGTGT 960
|||||
QY 961 CAGAGTCCGGCGGAACTTTTTCGAGCCAGGACCAACAACTGAAGACACAAA 1020
|||||
Db 961 CAGAGTCCGGCGGAACTTTTTCGAGCCAGGACCAACAACTGAAGACACAAA 1020
|||||
QY 1021 ATCATGCTTCAGAAAATTTCTTTCAGTGGTTCAGTGCACAGTCAAGGAGGAGCT 1080
|||||
Db 1021 ATCATGCTTCAGAAAATTTCTTTCAGTGGTTCAGTGCACAGTCAAGGAGGAGCT 1080
|||||
QY 1081 GCAGTGTGCAATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
|||||
Db 1081 GCAGTGTGCAATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
|||||
QY 1141 AAACAGGTCCGGACAAAGCACCCATTAATACACCCGTTGATAAACTTGACATCTCTGAG 1200
|||||
Db 1141 AAACAGGTCCGGACAAAGCACCCATTAATACACCCGTTGATAAACTTGACATCTCTGAG 1200
|||||
QY 1201 GCACTCAAGTTGAACACATCATCCGAGACAGACAGACAGACAGACAGACAGACAGCT 1260
|||||
Db 1201 GCACTCAAGTTGAACACATCATCCGAGACAGACAGACAGACAGACAGACAGACAGCT 1260
|||||
QY 1261 CCCTCTGCCAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCC 1320
|||||
Db 1261 CCCTCTGCCAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCC 1320
|||||
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCGAGACCGCTGGC 1380
|||||
Db 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCGAGACCGCTGGC 1380
|||||
QY 1381 AAGAAACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
|||||
Db 1381 AAGAAACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
|||||
QY 1441 TTAATTAACCAATCTATTTGCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
|||||
Db 1441 TTAATTAACCAATCTATTTGCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
|||||
QY 1501 CGATCG 1506
|||||
Db 1501 CGATCG 1506
|||||

```
; APPLICANT: WILSON, JULIE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND COMPLEMENTARITY-DETERMINING
; FILE REFERENCE: ARMY 166
; CURRENT APPLICATION NUMBER: US/10/226,795
; CURRENT FILING DATE: 2002-11-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola virus
US-10-226-795-1

Query Match      66.7%; Score 1501.4; DB 18; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTGCAATTGACGTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 60
DB 59 ATGGGCGTTACAGGAATATTGCAATTGACGTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 118
QY 61 CTTTGGGTAAATATCCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCCACAAT 120
DB 119 CTTTGGGTAAATATCCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCCACAAT 178
QY 121 AGCACATTACAGGTAGTGATGTCGACAAACTAGTTTGTGCGACAAACTGTCATCCACA 180
DB 179 AGCACATTACAGGTAGTGATGTCGACAAACTAGTTTGTGCGACAAACTGTCATCCACA 238
QY 181 AATCAATTGAGATCAGTTGGATCGAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 240
DB 239 AATCAATTGAGATCAGTTGGATCGAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 298
QY 241 TCTGCACTAAAGATGGGGCTTCAGGTCCGGTGTCCACCAAGGTGCTCAATATGAA 300
DB 299 TCTGCACTAAAGATGGGGCTTCAGGTCCGGTGTCCACCAAGGTGCTCAATATGAA 358
QY 301 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 360
DB 359 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 418
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGGCGGTATGTGCACAAA 420
DB 419 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGGCGGTATGTGCACAAA 478
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCCCTCCATAAAGAGGTGCTTTCTTC 480
DB 479 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCCCTCCATAAAGAGGTGCTTTCTTC 538
QY 481 CTGTATGATCGACTTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGCAAGGTGTC 540
DB 539 CTGTATGATCGACTTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGCAAGGTGTC 598
QY 541 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
DB 599 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 658
QY 601 GAGCCGGTCAATGCAACGGAGACCCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB 659 GAGCCGGTCAATGCAACGGAGACCCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 718
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTTCACAATTTGACC 720
DB 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTTCACAATTTGACC 778
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 779 TAGCTCCAACTTGAATCAAGATTTCACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 838
QY 781 TATACAAGTGGGAAAAGGAGCAATACACCGGGAAAACCTAATTTGGAGGTCAACCCCGAA 840
DB 839 TATACAAGTGGGAAAAGGAGCAATACACCGGGAAAACCTAATTTGGAGGTCAACCCCGAA 898
```

RESULT 4

```
US-10-491-121-6
; Sequence 6, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 6887
; TYPE: DNA

841 ATTGATACAAATCGGGAGTGGGCTTTCTGGAAACTAAAAAACTCCTCCTAGAAAA 900
899 ATTGATACAAATCGGGAGTGGGCTTTCTGGAAACTAAAAAACTCCTCCTAGAAAA 958
901 ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 960
959 ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 1018
961 CAGAGTCCGGCGCAATCTTCTTCGACCCAGGACCAACAACTGTAAGACCAACAA 1020
1019 CAGAGTCCGGCGCAATCTTCTTCGACCCAGGACCAACAACTGTAAGACCAACAA 1078
1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
1079 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1138
1081 GCAGTGTGCAATCTAAACAACTTTCGCAATCTCCACAGAGTCCCAATCCCTCAACCC 1140
1139 GCAGTGTGCAATCTAAACAACTTTCGCAATCTCCACAGAGTCCCAATCCCTCAACCC 1198
1141 AAACCGAGTCCGGACCAACAGCACCAATTAATACCCGTTGTAATACTTGACATCTCTGAG 1200
1199 AAACCGAGTCCGGACCAACAGCACCAATTAATACCCGTTGTAATACTTGACATCTCTGAG 1258
1201 GCAACTCAAGTTGAACAACTACCGCAGAAACAGACAAACAGACAGCAGCAGCCTCCGACACT 1260
1259 GCAACTCAAGTTGAACAACTACCGCAGAAACAGACAAACAGACAGCAGCAGCCTCCGACACT 1318
1261 CCTCTGCGACAGCGAGCGGACCCCAAGAGAGAGAACAGAGAGAACAGAGAGAGC 1320
1319 CCTCTGCGACAGCGAGCGGACCCCAAGAGAGAGAACAGAGAGAACAGAGAGAGC 1378
1321 ACTGACTTCTGGACCCCGCCACACACAACTGTCCTCCCAACACAGCAGAGACCGCTGCG 1380
1379 ACTGACTTCTGGACCCCGCCACACAACTGTCCTCCCAACACAGCAGAGAGAGCCTGCG 1438
1381 AACAACTCACTCATCAACAGATACCGGAGAGAGAGTGCAGCAGCGGAGAGCTAGGC 1440
1439 AACAACTCACTCATCAACAGATACCGGAGAGAGAGTGCAGCAGCGGAGAGCTAGGC 1498
1441 TTAATTACCAATATCTATTGCTGGAGTCCAGAGCTGATCAAGGCGGGAGAGAACTCGA 1500
1499 TTAATTACCAATATCTATTGCTGGAGTCCAGAGCTGATCAAGGCGGGAGAGAACTCGA 1558
1501 CGA 1503
1559 AGA 1561
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(Z) delta GP2 delta C-term A
US-10-491-121-6

Query Match          66.7%; Score 1501.4; DB 20; Length 6887;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTACAGGAATATTCAGTTACTCGTGTGATCAATCAAGAGACATCATTCCTTT 60
Db 1939 ATGGCGGTACAGGAATATTCAGTTACTCGTGTGATCAATCAAGAGACATCATTCCTTT 1998
QY 61 CTTTGGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 120
Db 1999 CTTTGGGTAAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 2058
QY 121 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 180
Db 2059 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTACAGTCCGGTGTCCCAACMAAGGTGTCATTTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGGCTTACAGTCCGGTGTCCCAACMAAGGTGTCATTTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAACTGCTTACAACTTGAATCAAAAAACCTGACGGAGTGCAG 360
Db 2239 GCTGTGTAATGGGCTGAAACTGCTTACAACTTGAATCAAAAAACCTGACGGAGTGCAG 2298
QY 361 TGTCTACAGCAGCCGACGAGCGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCCGACGAGCGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTCTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTCTC 2478
QY 541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCGGTCAATGCAACGGAGGACCGGTCTAGTGTCTACTATTTTACCACAATTAGATAT 660
Db 2539 GAGCGGTCAATGCAACGGAGGACCGGTCTAGTGTCTACTATTTTACCACAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTCAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTCAATGAGACAATA 2718
QY 781 TATCAAGTGGGAAAAGGAGCAATACCAAGGAAAACCTAAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATCAAGTGGGAAAAGGAGCAATACCAAGGAAAACCTAAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAA 2838
QY 901 ATTTCAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGAGCCAAAAACATCAGTGT 960
Db 2839 ATTTCAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGAGCCAAAAACATCAGTGT 2898
```

```
QY 961 CAGAGTCCGGCGGAACCTTCTTCGACCCAGGGACCAACAACTGAAGACCAAAA 1020
Db 2899 CAGAGTCCGGCGGAACCTTCTTCGACCCAGGGACCAACAACTGAAGACCAAAA 2958
QY 1021 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db 2959 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 3018
QY 1081 GCAGTGTGGANTTAACAACCCCTTGGCCAAATCTTCCAGAGTCCCAATCCCTCAACAAC 1140
Db 3019 GCAGTGTGGANTTAACAACCCCTTGGCCAAATCTTCCAGAGTCCCAATCCCTCAACAAC 3078
QY 1141 AAACCAAGTCCGGACACACAGCCCAATAATACACCCGTTATTAACCTTGACATCTCTGAG 1200
Db 3079 AAACCAAGTCCGGACACACAGCCCAATAATACACCCGTTATTAACCTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACCAACATCACCGCAGAACAGACAAACGACAGACAGCCCTCGACACT 1260
Db 3139 GCAACTCAAGTTGAACCAACATCACCGCAGAACAGACAAACGACAGACAGCCCTCGACACT 3198
QY 1261 CCCTCTGCCAGCACCGCAGCGGACCCCAAAAGCAGAGAACCAACACAGCAAGAGC 1320
Db 3199 CCCTCTGCCAGCACCGCAGCGGACCCCAAAAGCAGAGAACCAACACAGCAAGAGC 3258
QY 1321 ACTGACTTCTGGACCCGCGCCACCAACAGTCCCAAAACACAGCGAGACCGCTGGC 1380
Db 3259 ACTGACTTCTGGACCCGCGCCACCAACAGTCCCAAAACACAGCGAGACCGCTGGC 3318
QY 1381 AACAAACACTCATATCAACAGATACCGGAGAGAGTGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAAACACTCATATCAACAGATACCGGAGAGAGTGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATATCTATTTGAGTCCGAGGACTGATCAAGCGGGAGAGAACTCGA 1500
Db 3379 TTAATTACCAATATCTATTTGAGTCCGAGGACTGATCAAGCGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441

RESULT 5
US-10-491-121-9
; Sequence 9, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ. ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6914
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(Z) delta TM
US-10-491-121-9

Query Match          66.7%; Score 1501.4; DB 20; Length 6914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGTACAGGAATATTGCAAGTTACCTCGTGAATCGAATCAAGAGGACATCATCTTTT 60
Db 1939 ATGGCGGTACAGGAATATTGCAAGTTACCTCGTGAATCGAATCAAGAGGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 120
Db 1999 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 2058
QY 121 AGCACATTACAGGTTAGTGTGCGACAACTAGTGTGCGTGAACAACTGTCATCCACA 180
Db 2059 AGCACATTACAGGTTAGTGTGCGACAACTAGTGTGCGTGAACAACTGTCATCCACA 2118
QY 181 AATCAATTAGATCAGTTGGATCGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCA 240
Db 2119 AATCAATTAGATCAGTTGGATCGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTCAAGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGGCTTCAAGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 2238
QY 301 GCTGTGTAATGGGCTGMAAACTGCTACAACTTTGAAATCAAAAACTTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGMAAACTGCTACAACTTTGAAATCAAAAACTTGACGGGAGTGAG 2298
QY 361 TGTCTACCAAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACCAAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 2478
QY 541 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAATTTAGATAT 660
Db 2539 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAATTTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 2658
QY 721 TAGTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAGTGGGAAAGGAGCAATACCACGGGAAACTTAATTTGGAAGTCAACCCGAA 840
Db 2719 TATACAGTGGGAAAGGAGCAATACCACGGGAAACTTAATTTGGAAGTCAACCCGAA 2778
QY 841 ATTGATACAACTACGGGAGTGGGCTTCTCGGAACTTAAAGAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTACGGGAGTGGGCTTCTCGGAACTTAAAGAAACCTCACTAGAAAA 2838
QY 901 ATTTCGAGTGAAGAGTTGTCTTTTCAGTTGTATCAAAACGGAGCCCAAAACATCATGTGT 960
Db 2839 ATTTCGAGTGAAGAGTTGTCTTTTCAGTTGTATCAAAACGGAGCCCAAAACATCATGTGT 2898
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGAACCAACAACTGAAAGCACACAAA 1020
Db 2899 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGAACCAACAACTGAAAGCACACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGGAAGCT 3018
QY 1081 GCAGTGTGCAATCTAAACAAACCTTGCACAAATCTCCACGAGTCCCAATCCCTCAACACC 1140
```

```
Db 3019 GCAGTGTGCAATCTAAACAACCTTTGCCACAATCTCCACGAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACAGGTCCGGACAAACAGCACCAATAATACACCGTGTATAAACTTGAACATCTCTGAG 1200
Db 3079 AAACAGGTCCGGACAAACAGCACCAATAATACACCGTGTATAAACTTGAACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACTATCACCGCAGAAACAGACAAACAGACAGCTTCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACTATCACCGCAGAAACAGACAAACAGACAGCTTCGACACT 3198
QY 1261 CCCTCTGCCAGCAGCGAGCCGGACCCCAAAAGCAGAGAAACCAACACGAGCAAGAGC 1320
Db 3199 CCCTCTGCCAGCAGCGAGCCGGACCCCAAAAGCAGAGAAACCAACACGAGCAAGAGC 3258
QY 1321 ACTGACTTCTTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCGAGACCCCTGGC 1380
Db 3259 ACTGACTTCTTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCGAGACCCCTGGC 3318
QY 1381 AACAAACACTCATCAACAAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAAACACTCATCAACAAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCACAGCGGGAGAGACTCGA 1500
Db 3379 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCACAGCGGGAGAGACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
```

RESULT 6

```
US-10-491-121-7
; Sequence 7, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 7044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(Z) delta GP2 Delta C-term B
US-10-491-121-7
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 7044;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGTACAGGAATATTGCAAGTTACCTCGTGAATCGAATCAAGAGGACATCATCTTTT 60
Db 1939 ATGGCGGTACAGGAATATTGCAAGTTACCTCGTGAATCGAATCAAGAGGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 120
Db 1999 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 2058
QY 121 AGCACATTACAGGTTAGTGTGCGACAAACTAGTTTTGTGCGTGAACAACTGTCATCCACA 180
```

```
Db 2059 AGCAATTACAGTTAGTGTGATGTCGACAAACTAGTGTGTCGACAAACTGTCTATCCACA 2118
Qy 181 AATCAATTGAGATCAGTTGGATCTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGATCTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 2178
Qy 241 TCTGCACTAAAGATGGGGCTTCAAGTCCGGTGTCCCAACAAAGGTGGTCAATATATGAA 300
Db 2179 TCTGCACTAAAGATGGGGCTTCAAGTCCGGTGTCCCAACAAAGGTGGTCAATATATGAA 2238
Qy 301 GCTGTGATGGGCTGAAACTGCTTACAACTCTTGAATCAAAACCTGACGGAGTGAG 360
Db 2239 GCTGTGATGGGCTGAAACTGCTTACAACTCTTGAATCAAAACCTGACGGAGTGAG 2298
Qy 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTTCCATAAAGGGTCTTCTTC 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTTCCATAAAGGGTCTTCTTC 2358
Qy 421 GTATCAGGAA CGGACCGTGTGCGGAGACTTTGCTTCCATAAAGGGTCTTCTTC 480
Db 2359 GTATCAGGAA CGGACCGTGTGCGGAGACTTTGCTTCCATAAAGGGTCTTCTTC 2418
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGAACGACTTTCGCTGAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGAACGACTTTCGCTGAGGTGTC 2478
Qy 541 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
Qy 601 GAGCCGGTCAATGCAACGGAGACCCGCTCTAGTGGCTACTATTTACACAAATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGGAGACCCGCTCTAGTGGCTACTATTTACACAAATTAGATAT 2598
Qy 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGTTTGACAAATTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGTTTGACAAATTGACC 2658
Qy 721 TACGTCCTCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TACGTCCTCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
Qy 781 TATACAAGTGGGAAAGAGGCAATACCAACGGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAAGAGGCAATACCAACGGGAAACTAAATTTGGAAGGTCAACCCCGAA 2778
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAAACCTCACTAGAAAA 2838
Qy 901 ATTCGAGTGAAGATTGTCTTTTACAGTTGTATCAACCGGAGCCCAAAAAATCATCAGTGGT 960
Db 2839 ATTCGAGTGAAGATTGTCTTTTACAGTTGTATCAACCGGAGCCCAAAAAATCATCAGTGGT 2898
Qy 961 CAGAGTCCGGCCGAACTTCTTCCGACCCGAGGACCAACACAACTCTGAAGACCAACAA 1020
Db 2899 CAGAGTCCGGCCGAACTTCTTCCGACCCGAGGACCAACACAACTCTGAAGACCAACAA 2958
Qy 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGCTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAATTCCTCTGCAATGCTTCAAGTGCACAGTCAAGGAAGGAAGCT 3018
Qy 1081 GCAGTGTGCACTCTAACAACTTCCCACTCTCCAGAGTCCCAATCTCCAGAGTCCCACTCAAC 1140
Db 3019 GCAGTGTGCACTCTAACAACTTCCCACTCTCCAGAGTCCCAATCTCCAGAGTCCCACTCAAC 3078
Qy 1141 AAACCAAGTCCCGACAAACAGCACCATAATACACCCGCTGTATATACTTGACATCTCTGAG 1200
Db 3079 AAACCAAGTCCCGACAAACAGCACCATAATACACCCGCTGTATATACTTGACATCTCTGAG 3138
Qy 1201 GCAACTCAAGTTGAACACATCACCGCAGACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 1260
```

```
Db 3139 GCAACTCAAGTTGAACAACTATCCCGCAGAACAGACAAACAGCAGCAGCCTCCGACACT 3198
Qy 1261 CCCTCTCCAGCAGCCGAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 1320
Db 3199 CCCTCTCCAGCAGCCGAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 3258
Qy 1321 ACTGACTTCTGGACCCGCGCCACCAACAACTCCCAAAACCAACAGGAGACCGCTGGC 1380
Db 3259 ACTGACTTCTGGACCCGCGCCACCAACAACTCCCAAAACCAACAGGAGACCGCTGGC 3318
Qy 1381 AACAAACAACTCATCAACCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAAACAACTCATCAACCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 3378
Qy 1441 TTAATTACCAATACTATTGTGGAGTCCGAGGACTGATCAACAGCGGAGAGAACTCGA 1500
Db 3379 TTAATTACCAATACTATTGTGGAGTCCGAGGACTGATCAACAGCGGAGAGAACTCGA 3438
Qy 1501 CGA 1503
Db 3439 AGA 3441

RESULT 7
US-10-491-121-8
; Sequence 8, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 7106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(2) delta GP2 delta FUS
US-10-491-121-8

Query Match 66.7%; Score 1501.4; DB 20; Length 7106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGTTACAGGAATATTTCAGTTACCTCGTGTATCGATTCAAGAGACATCATTTCTTT 60
Db 1939 ATGGGCGTTACAGGAATATTTCAGTTACCTCGTGTATCGATTCAAGAGACATCATTTCTTT 1998
Qy 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
Db 1999 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 2058
Qy 121 AGCACTTTACAGGTTAGTGTGTCGACAAACTAGTTTGTGTCGACAAACTGTCTATCCACA 180
Db 2059 AGCACTTTACAGGTTAGTGTGTCGACAAACTAGTTTGTGTCGACAAACTGTCTATCCACA 2118
Qy 181 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 2178
Qy 241 TCTGCACTAAAGATGGGGCTTCAAGTCCGGTGTCCCAACAAAGGTGGTCAATATATGAA 300
```



```
Db 2179 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCGCCACCAAGAGGTGCTCAATTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGGTATGTGACAAA 420
Db 2299 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGGTATGTGACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGCGCGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGCGCGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGNAACGACTTTGCTGAAGGTGTC 2478
QY 541 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACCGAGGACCGTCTAGTGGCTACTATTCTACCAANTTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACCGAGGACCGTCTAGTGGCTACTATTCTACCAANTTAGATAT 2598
QY 661 CAGGCTACCGGTTTTGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTTGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
QY 721 TACGTCCAACTTGAATCAAGATTTCACACAGTTTCTGCTCAGCTGAATGAGACAATA 780
Db 2659 TACGTCCAACTTGAATCAAGATTTCACACAGTTTCTGCTCAGCTGAATGAGACAATA 2718
QY 781 TATACAGTGGGAAGAGGAGCAATACCGAGGAACAAATTTGGAGTCAACCCCGAA 840
Db 2719 TATACAGTGGGAAGAGGAGCAATACCGAGGAACAAATTTGGAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 2838
QY 901 ATTGCAAGTGAAGATGTCTTTTCAAGTTGATCAACCGGACCAAAAAACATCAGTGT 960
Db 2839 ATTGCAAGTGAAGATGTCTTTTCAAGTTGATCAACCGGACCAAAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGGAGTCTTCCGACCCAGGACCAACACACACTGAAGACCAACAA 1020
Db 2899 CAGAGTCCGGCGGAGTCTTCCGACCCAGGACCAACACACACTGAAGACCAACAA 2958
QY 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 3018
QY 1081 GCAGTGTGCGATCTAACAACCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCGATCTAACAACCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACGAGTCCGGACACAGACCCCAATATACCCGCTGTATAAATCTGACATCTCTGAG 1200
Db 3079 AAACGAGTCCGGACACAGACCCCAATATACCCGCTGTATAAATCTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAAATCACCGCAGAACAGACAGACAGACGCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAAATCACCGCAGAACAGACAGACAGACGCTCCGACACT 3198
QY 1261 CCTCTGCAACGACCGGACCGGACCCCAAGAGAGAGAAACCAACACGAGCAGAGAGC 1320
Db 3199 CCTCTGCAACGACCGGACCGGACCCCAAGAGAGAGAAACCAACACGAGCAGAGAGC 3258
QY 1321 ACTGACTTCTTGACCCCGGACCCCAACAGTCCCAAAACCAACGAGCAGACCCCTGCG 1380
Db 3259 ACTGACTTCTTGACCCCGGACCCCAACAGTCCCAAAACCAACGAGCAGACCCCTGCG 3318
```

RESULT 8

```
US-10-491-121-1
; Sequence 1, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7154
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 -GP(Z)
US-10-491-121-1
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 7154;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60
Db 1939 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 1998
QY 61 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCAATCCACAAT 120
Db 1999 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCAATCCACAAT 2058
QY 121 AGCATTACAGGTTAGTGATGTGCAAAAATAGTTTGTGTGACAAAATGTCTATCCACA 180
Db 2059 AGCATTACAGGTTAGTGATGTGCAAAAATAGTTTGTGTGACAAAATGTCTATCCACA 2118
QY 181 AATCAATTTGAGATCAGTTGGACTGAACTCTCGAAGGGAATGGAGTGGCACTGACGTGCCA 240
Db 2119 AATCAATTTGAGATCAGTTGGACTGAACTCTCGAAGGGAATGGAGTGGCACTGACGTGCCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTTCAGGTCGGGTGTCCCAACCAAGGTTGGTCAATTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGGCTTTCAGGTCGGGTGTCCCAACCAAGGTTGGTCAATTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAAACTGCTTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAAACTGCTTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCCGACAGAGCGGATTCGGGGCTTCCCGGTGCGGGTATGTGACAAAA 420
Db 2299 TGTCTACAGCAGCCGACAGAGCGGATTCGGGGCTTCCCGGTGCGGGTATGTGACAAAA 2358
```


QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCGCTTCCATAAAGAGGGTCTTCTTC 480
DB |||||
QY 481 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
DB |||||
QY 541 GTTGCAATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCACAACCCCTTGAGA 600
DB |||||
QY 601 GAGCCGGTCAATGCAACGGAGGACCGCTCTAGTGGCTACTATTCTACCAAAATTAGATAT 660
DB |||||
QY 661 CAGGCTACCGGTTTTGGAAACCAATCAGACAGAGTACTTTGTCGAGTTGACAATTTGACC 720
DB |||||
QY 721 TACGTCCTCACTTGAATCAAGATTTCAACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 780
DB |||||
QY 781 TATACAAGTGGGAAAGAGGCAATACACGGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
DB |||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA 900
DB |||||
QY 901 ATTCGAGTGAAGAGTGTCTTTTCAAGTTGATCAAAACGGAGCCAAAAACATCAGTGGT 960
DB |||||
QY 961 CAGAGTCCGGCGGAACTTCTTCCGACCGAGGACCAACAACTGTAAGACACACAAA 1020
DB |||||
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCAAGTCAAGGAGGGAAGCT 1080
DB |||||
QY 1081 GCAGTGTGCACTTAACAACCTTGCCCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
DB |||||
QY 1141 AAACGAGTCCGGACAACAGCAACCAATATACACCGGTGTATAAATTGACATCTCTGAG 1200
DB |||||
QY 1201 GCAACTCAAGTTGAACACATCACCGCAGACAGACAGACAGACAGACAGACAGACAG 1260
DB |||||
QY 1261 CCCTCTGCCACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 1320
DB |||||
QY 1321 ACTGACTTCTTGAGCCCGGCCACCAACAAGTCTCCCAAAACCAACGAGGAGCCGCTGGC 1380
DB |||||
QY 1381 AACAAACACTCATCACAAGATACCGGAGGAGGAGTCCAGGAGGAGGAGTCCAGGAGGAG 1440
DB |||||
QY 1441 TTAATTACCAATACTATTCTGAGTCCGAGGAGTATCAGAGGAGGAGGAGGAGGAGGAG 1500
DB |||||

QY 1501 CGA 1503
DB 3439 AGA 3441

RESULT 9
US-10-491-121-2
; Sequence 2, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7188
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Ebola GP(Z)
US-10-491-121-2

Query Match 66.7%; Score 1501.4; DB 20; Length 7188;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTGCACTTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 60
DB 1939 ATGGGCGTTACAGGAATATTGCACTTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 1998
QY 61 CTTTGGGTAATTAATCCCTTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCATCCACAAT 120
DB 1999 CTTTGGGTAATTAATCCCTTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCATCCACAAT 2058
QY 121 AGCATTACAGGTTAGTGTGTCGACAAACTAGTGTGTCGACAAACTGTCTCATCCACA 180
DB 2059 AGCATTACAGGTTAGTGTGTCGACAAACTAGTGTGTCGACAAACTGTCTCATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGAATCTGAAAGGGAATGAGTGGCAACTGAGCGTGCCA 240
DB 2119 AATCAATTGAGATCAGTTGGAATCTGAAAGGGAATGAGTGGCAACTGAGCGTGCCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAGAGGTGGTCAATTATGAA 300
DB 2179 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAGAGGTGGTCAATTATGAA 2238
QY 301 GCTGTTGAATGGGCTGAAAACCTGCTACAACTTCAAAATCAAAAACTGAGCGGAGTGAG 360
DB 2239 GCTGTTGAATGGGCTGAAAACCTGCTACAACTTCAAAATCAAAAACTGAGCGGAGTGAG 2298
QY 361 TGTCTACAGCAGCCGAGCGGATTCGGGGCTTCCCGGTCGGGTGTCGAGGATGTGTCACAAA 420
DB 2299 TGTCTACAGCAGCCGAGCGGATTCGGGGCTTCCCGGTCGGGTGTCGAGGATGTGTCACAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTTCCTTCCATAAAGAGGGTCTTCTTC 480
DB 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTTCCTTCCATAAAGAGGGTCTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAGGAGTCTTTCGCTGAAGGTGTC 540
DB 2419 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAGGAGTCTTTCGCTGAAGGTGTC 2478

```
QY 541 GTTGCAATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGAGGACCGCTAGTGCTACTATTCTACCAATATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGAGGACCGCTAGTGCTACTATTCTACCAATATTAGATAT 2598
QY 661 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTGGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTGGACAAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTACACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTACACACAGTTTCTGCTCCAGTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAGGAGCAATACACCGGAAACTAAATTTGGAAGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAGGAGCAATACACCGGAAACTAAATTTGGAAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAAACCTCACTAGAAAA 2838
QY 901 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGAGCCAAACATCATGTGT 960
Db 2839 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACGGAGCCAAACATCATGTGT 2898
QY 961 CAGAGTCCGGCGGACTTCTCCGACCCAGGACCCACACAACTGAAGACACACAAA 1020
Db 2899 CAGAGTCCGGCGGAACTTCTCCGACCCAGGACCCACACAACTGAAGACACACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGAAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGAAAGGGAAGCT 3018
QY 1081 GCAGTGTCCGACTCTAAACACCTTTGCCAACAATCTCCAGAGTCCCAATCCCTCAACAC 1140
Db 3019 GCAGTGTCCGACTCTAAACACCTTTGCCAACAATCTCCAGAGTCCCAATCCCTCAACAC 3078
QY 1141 AACCAAGTCCGGACACAGCACCCCATATACCCGGTGTATAAATTTGACATCTCTGAG 1200
Db 3079 AACCAAGTCCGGACACAGCACCCCATATACCCGGTGTATAAATTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGACAGCCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGCCTCCGACACT 3198
QY 1261 CCTCTGCCACGACCGCAGCCGGAACCCCAAGAGAGAGTCCAGACGAGGGAAGCTAGGC 1320
Db 3199 CCTCTGCCACGACCGCAGCCGGAACCCCAAGAGAGAGTCCAGACGAGGGAAGCTAGGC 3258
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCCAAACACAGCGAGACCCCTGGC 1380
Db 3259 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCCAAACACAGCGAGACCCCTGGC 3318
QY 1381 AACCAACAACATCATCAACAAGATACCGGAGAGAGTCCAGACGAGGGAAGCTAGGC 1440
Db 3319 AACCAACAACATCATCAACAAGATACCGGAGAGAGTCCAGACGAGGGAAGCTAGGC 3378
QY 1441 TTAATTAACAATATCTATTCTGGAGTCGAGACTGATCACAGGGGGAGAGAACTCGA 1500
Db 3379 TTAATTAACAATATCTATTCTGGAGTCGAGACTGATCACAGGGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
```

RESULT 10
US-10-860-878-3
; Sequence 3, Application US/10860878
; Publication No. US20050130129A1

```
; GENERAL INFORMATION:  
; APPLICANT: Nabel, Gary J  
; APPLICANT: Delgado, Rafael  
; APPLICANT: Yang, Zhi-yong  
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types  
; TITLE OF INVENTION: By Pseudotyping With Viral Glycoprotein  
; FILE REFERENCE: 21159-001474POA  
; CURRENT APPLICATION NUMBER: US/10/860,878  
; PRIOR FILING DATE: 2004-06-03  
; PRIOR APPLICATION NUMBER: US/09/600,766  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: USSN 60/072033  
; PRIOR FILING DATE: 1998-01-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 7285  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid  
; OTHER INFORMATION: containing DNA of GP of Ebola Virus, Zaire strain  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (886)..(1129)  
; OTHER INFORMATION: CMV IE 5' UT  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1130)..(1840)  
; OTHER INFORMATION: CMV IE INT  
; FEATURE:  
; NAME/KEY: enhancer  
; LOCATION: (248)..(885)  
; OTHER INFORMATION: CMV enhancer  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1870)..(4301)  
; OTHER INFORMATION: GP (Z)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (4302)..(4854)  
; OTHER INFORMATION: TBGH  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (6350)..(6972)  
; OTHER INFORMATION: Kan r  
; US-10-860-878-3
```

```
Query Match 66.7%; Score 1501.4; DB 22; Length 7285;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 60  
Db 2072 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 2131  
QY 61 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120  
Db 2132 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 2191  
QY 131 AGCAATTACAGGTAGTGTGACAAACTAGTTTTCGTGACAAACTGTGTATCCACA 180  
Db 2192 AGCAATTACAGGTAGTGTGACAAACTAGTTTTCGTGACAAACTGTGTATCCACA 2251  
QY 181 AATCAATTTGAGTTCAGTTGGACTGAATCTCGAAGGAATGGAGTGGCACTGACCTGCCA 240  
Db 2252 AATCAATTTGAGTTCAGTTGGACTGAATCTCGAAGGAATGGAGTGGCACTGACCTGCCA 2311  
QY 241 TCTGCAACTAAAGATGGGGCTTTCAGGTCGCGGTGCCCAACCAAGAGGTGGTCAATTTATGAA 300  
Db 2312 TCTGCAACTAAAGATGGGGCTTTCAGGTCGCGGTGCCCAACCAAGAGGTGGTCAATTTATGAA 2371  
QY 301 GCTGTGTAATGGGCTGAAAACTGCTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
```

Db 2372 GCTGTGTAATGGCTGAAAACTGCTCAATCTTGAATCAAAAAACCTTGACGGGAGTGAG 2431
Qy 361 TGTCTACAGCAGCCGACAGCGGATTCGGGCTTTCCCGGTGCGGTATGTGCACAAA 420
Db 2432 TGTCTACAGCAGCCGACAGCGGATTCGGGCTTTCCCGGTGCGGTATGTGCACAAA 2491
Qy 421 GTATCAGGACGGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2492 GTATCAGGACGGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTCTTTCTTC 2551
Qy 481 CTGTATGATCGACTTTCCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 2552 CTGTATGATCGACTTTCCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 2611
Qy 541 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2612 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2671
Qy 601 GAGCGGTCAATGCAACCGGAGGACCGGTCTAGTGGCTACTATTCTACCAATTAGATAT 660
Db 2672 GAGCGGTCAATGCAACCGGAGGACCGGTCTAGTGGCTACTATTCTACCAATTAGATAT 2731
Qy 661 CAGGCTACCGGTTTTGGGAACCAATCAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 2732 CAGGCTACCGGTTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 2791
Qy 721 TACGTCCTCAATTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2792 TACGTCCTCAATTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2851
Qy 781 TATACAGTGGGAAAAAGAGGACATACACGGGAAAACTAATTTGGAAGGTCAACCCCGAA 840
Db 2852 TATACAGTGGGAAAAAGAGGACATACACGGGAAAACTAATTTGGAAGGTCAACCCCGAA 2911
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAACTCACTAGAGAAA 900
Db 2912 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAACTCACTAGAGAAA 2971
Qy 901 ATTCGAGTGAAGAGTTGTCTTTTCACAGTTGTATCAAAACGGGACCAAAAACATCAGTGGT 960
Db 2972 ATTCGAGTGAAGAGTTGTCTTTTCACAGTTGTATCAAAACGGGACCAAAAACATCAGTGGT 3031
Qy 961 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGACCAACAACTCTGAAGACCAACAAA 1020
Db 3032 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGACCAACAACTCTGAAGACCAACAAA 3091
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCAAGTCAAGGAGGAGGAGCT 1080
Db 3092 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCAAGTCAAGGAGGAGGAGCT 3151
Qy 1081 GCAGTGTGCGATCTTAACAACTTCCCAATCTCCACGAGTCCCAATCTCCCAAC 1140
Db 3152 GCAGTGTGCGATCTTAACAACTTCCCAATCTCCACGAGTCCCAATCTCCCAAC 3211
Qy 1141 AAACGAGTCCGGACAAAGCAACCAATATAACCCGCTGTATAAATTGACATCTCTGAG 1200
Db 3212 AAACGAGTCCGGACAAAGCAACCAATATAACCCGCTGTATAAATTGACATCTCTGAG 3271
Qy 1201 GCAACTCAAGTTGAACACATCACCGGAGACAGACACGACGACGACGCTCCGACACT 1260
Db 3272 GCAACTCAAGTTGAACACATCACCGGAGACAGACACGACGACGACGCTCCGACACT 3331
Qy 1261 CCCTCTGCGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 1320
Db 3332 CCCTCTGCGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 3391
Qy 1321 ACTGATCTCTGAGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 1380
Db 3392 ACTGATCTCTGAGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 3451
Qy 1381 AACAACACACTCATCAACCAATACCGGAGAGAGTGGCCAGCAGCGGGAGCTAGGC 1440

Db 3452 AACAAACAACTCATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGAGGACTAGGC 3511
Qy 1441 TTAATTAACCAATATATTTCTGGAGTGCAGGACTGATCAACAGCGGAGAGAACTCGA 1500
Db 3512 TTAATTAACCAATATATTTCTGGAGTGCAGGACTGATCAACAGCGGAGAGAACTCGA 3571
Qy 1501 CGA 1503
Db 3572 AGA 3574
RESULT 11
US-10-491-121-29
; Sequence 29, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 8199
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Ebola GP(2) (dtm)
US-10-491-121-29
Query Match 66.7%; Score 1501.4; DB 20; Length 8199;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATATTGCACTTACCTCGTGATCGAATTCAGAGAGACATCATCTTTT 60
Db 1336 ATGGGCGTTACAGGAATATTGCACTTACCTCGTGATCGAATTCAGAGAGACATCATCTTTT 1395
Qy 61 CTTTGGGTTAATATCTCTTTCCAAAGAGACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 1396 CTTTGGGTTAATATCTCTTTCCAAAGAGACATTTTCCATCCCACTTGGAGTCATCCACAAT 1455
Qy 121 AGCACATTACAGGTTAGTGATGTGCAACAACTAGTTTGTCTGTGACAAACTGTCTATCCACA 180
Db 1456 AGCACATTACAGGTTAGTGATGTGCAACAACTAGTTTGTCTGTGACAAACTGTCTATCCACA 1515
Qy 181 AATCAATTGAGATCAGTTTGGACTGAATCTCGAAGGGAATGAGATGGCAACTGACGTGCA 240
Db 1516 AATCAATTGAGATCAGTTTGGACTGAATCTCGAAGGGAATGAGATGGCAACTGACGTGCA 1575
Qy 241 TCTCNACTAAAGATGGGCTTTCAGGTCGGGTGTCCCAACCAAGGTGCTCAATTTATGAA 300
Db 1576 TCTCNACTAAAGATGGGCTTTCAGGTCGGGTGTCCCAACCAAGGTGCTCAATTTATGAA 1635
Qy 301 GCTGTGTAATGGGCTGAAAACTGCTTACAATCTTGAAATCAAAAAACCTTGACGGGAGTGAG 360
Db 1636 GCTGTGTAATGGGCTGAAAACTGCTTACAATCTTGAAATCAAAAAACCTTGACGGGAGTGAG 1695
Qy 361 TGTCTACAGCAGCCGACGAGATTCGGGCTTTCGCCGCTGCGGTATGTGCACAAA 420
Db 1696 TGTCTACAGCAGCCGACGAGATTCGGGCTTTCGCCGCTGCGGTATGTGCACAAA 1755
Qy 421 GTATCAGGAAACCGGACCGGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTCTTTCTTC 480

```
Db 1756 GTATCAGGAACGGGACCGGTGCGGAGAGACTTTGCTTCCATAAAGAGGGTGTCTTCTTC 1815
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db 1816 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 1875
Qy 541 GTTGCATTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCCAGCTCACACCCCTTGAGA 600
Db 1876 GTTGCATTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCCAGCTCACACCCCTTGAGA 1935
Qy 601 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGTCTACTATTCTACCAATATAGATAT 660
Db 1936 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGTCTACTATTCTACCAATATAGATAT 1995
Qy 661 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGACAAATTTGACC 720
Db 1996 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGACAAATTTGACC 2055
Qy 721 TACGTCCTCAACTTGAATCAAGATTCTACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2056 TACGTCCTCAACTTGAATCAAGATTCTACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2115
Qy 781 TATACAAGTGGGAAAGAGGCAATACACCGGGAATACTAATTTGGAAGTCAACCCGAA 840
Db 2116 TATACAAGTGGGAAAGAGGCAATACACCGGGAATACTAATTTGGAAGTCAACCCGAA 2175
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAACTTAAAGAACTTAAAGAA 900
Db 2176 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAACTTAAAGAACTTAAAGAA 2235
Qy 901 ATTCCAGTGAAGAGTGTCTTTCTACAGTTGATCAACCGGAGCCAAACCAATCATCAGTGT 960
Db 2236 ATTCCAGTGAAGAGTGTCTTTCTACAGTTGATCAACCGGAGCCAAACCAATCATCAGTGT 2295
Qy 961 CAGAGTCCGGCCGGAATCTTCTCCAGCCAGGAGCAACACAACTGAAGACACACAA 1020
Db 2296 CAGAGTCCGGCCGGAATCTTCTCCAGCCAGGAGCAACACAACTGAAGACACACAA 2355
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 2356 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 2415
Qy 1081 GCAGTGTGCACTTAACAAACCTTGCACAACTCTCCAGAGTCCCAATCTCCCTCAACACC 1140
Db 2416 GCAGTGTGCACTTAACAAACCTTGCACAACTCTCCAGAGTCCCAATCTCCCTCAACACC 2475
Qy 1141 AAACAGGTCCGGACACAGACCCCATATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 2476 AAACAGGTCCGGACACAGACCCCATATATACACCGGTGTATAAATTGACATCTCTGAG 2535
Qy 1201 GCAACTCAAGTTGAACCAATCAACGAGAACAGACAGACAGCAGCCTCCGACACT 1260
Db 2536 GCAACTCAAGTTGAACCAATCAACGAGAACAGACAGACAGCAGCCTCCGACACT 2595
Qy 1261 CCCTCTGCCACGACCGGACCGGACCCCAAGAGAGAGAACCAACACGAGCAGGAGAGC 1320
Db 2596 CCCTCTGCCACGACCGGACCGGACCCCAAGAGAGAGAACCAACACGAGCAGGAGAGC 2655
Qy 1321 ACTGACTTCTGGACCCGCCACCAACAACTGCCCAAAACCAACGAGAGACCCCTGGC 1380
Db 2656 ACTGACTTCTGGACCCGCCACCAACAACTGCCCAAAACCAACGAGAGACCCCTGGC 2715
Qy 1381 AACAAACACACTCATCACCAGATACCGGAGAGAGAGTCCAGACGCGGGAAGCTAGGC 1440
Db 2716 AACAAACACACTCATCACCAGATACCGGAGAGAGAGTCCAGACGCGGGAAGCTAGGC 2775
Qy 1441 TTAATTACCAATATCTATTCTGGAGTCGAGGATGATCACAGGCGGGAAGAACTCGA 1500
Db 2776 TTAATTACCAATATCTATTCTGGAGTCGAGGATGATCACAGGCGGGAAGAACTCGA 2835
Qy 1501 CGA 1503
Db 2836 AGA 2838
```

RESULT 12

```
US-10-491-121-28
; Sequence 28, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 8439
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pAdapt Ebola GP(Z)
US-10-491-121-28
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 8439;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGGTACAGGAATATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTT 60
Db 1336 ATGGCGGTACAGGAATATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTT 1395
Qy 61 CTTTGGGTAATATTCCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 1396 CTTTGGGTAATATTCCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCATCCACAAT 1455
Qy 121 AGCATTATACAGGTAGTGTGTGACAAACTAGTTTCTGTGACAAACTGTGTATCCACA 180
Db 1456 AGCATTATACAGGTAGTGTGTGACAAACTAGTTTCTGTGACAAACTGTGTATCCACA 1515
Qy 181 AATCAATTTGATCAGTTGGACTGNACTCCAGGGAATGGAGTGGCACTGACCTGCCA 240
Db 1516 AATCAATTTGATCAGTTGGACTGNACTCCAGGGAATGGAGTGGCACTGACCTGCCA 1575
Qy 241 TCTGCAACTAAAGATGGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATATGAA 300
Db 1576 TCTGCAACTAAAGATGGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATATGAA 1635
Qy 301 GCTGTGAATGGGTGAAAACTGTGTACAACTTTGAAATCAAAAAACCTTGACGGGAGTGAG 360
Db 1636 GCTGTGAATGGGTGAAAACTGTGTACAACTTTGAAATCAAAAAACCTTGACGGGAGTGAG 1695
Qy 361 TGTCTACAGCAGCGCCAGCGGATTCGGGCTTCCCGGTGCGCGGTATGTGACACAA 420
Db 1696 TGTCTACAGCAGCGCCAGCGGATTCGGGCTTCCCGGTGCGCGGTATGTGACACAA 1755
Qy 421 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 1756 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 1815
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db 1816 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 1875
Qy 541 GTTGCATTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCCAGCTCACACCCCTTGAGA 600
Db 1876 GTTGCATTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCCAGCTCACACCCCTTGAGA 1935
```

```
QY 601 GAGCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATTCTTACCACAAATTAGATAT 660
Db |||||
QY 1936 GAGCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATTCTTACCACAAATTAGATAT 1995
Db |||||
QY 661 CAGGCTACCGGTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTTGACAATTGACC 720
Db |||||
QY 1996 CAGGCTACCGGTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTTGACAATTGACC 2055
Db |||||
QY 721 TACGTCCAACTTGAATCAAGATTTCACACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 780
Db |||||
QY 2056 TACGTCCAACTTGAATCAAGATTTCACACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 2115
Db |||||
QY 781 TATACAAGTGGGAAAGAGCAATACACCGGGAACCTAATTGGAAGTCAACCCCGAA 840
Db |||||
QY 2116 TATACAAGTGGGAAAGAGCAATACACCGGGAACCTAATTGGAAGTCAACCCCGAA 2175
Db |||||
QY 841 ATTGATACAACTCGGGAGTGGGCCCTTCTGGGAACTTAAATAAACTCACTAGAAAA 900
Db |||||
QY 2176 ATTGATACAACTCGGGAGTGGGCCCTTCTGGGAACTTAAATAAACTCACTAGAAAA 2235
Db |||||
QY 901 ATTCGAGTGAAGATTGTCTTTTACAGTTGTATCAAAACGGACCCAAACATCAGTGGT 960
Db |||||
QY 2236 ATTCGAGTGAAGATTGTCTTTTACAGTTGTATCAAAACGGACCCAAACATCAGTGGT 2295
Db |||||
QY 961 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAACTTGAAGACCAACAA 1020
Db |||||
QY 2296 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAACTTGAAGACCAACAA 2355
Db |||||
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db |||||
QY 2356 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 2415
Db |||||
QY 1081 GCAGTGTGGCTCTTAACAACTTCCCAATCTCCAGAGTCCCAATCTCCACAGTCCCAATCT 1140
Db |||||
QY 2416 GCAGTGTGGCTCTTAACAACTTCCCAATCTCCAGAGTCCCAATCTCCACAGTCCCAATCT 2475
Db |||||
QY 1141 AAACAGGTCCGGACAAACAGCAACCCATTAATACACCCGTGTATAAATTGACATCTCTGAG 1200
Db |||||
QY 2476 AAACAGGTCCGGACAAACAGCAACCCATTAATACACCCGTGTATAAATTGACATCTCTGAG 2535
Db |||||
QY 1201 GCAACTCAAGTTGAACATCATCCGACAGACAGACAGACAGACAGACAGACAGACAGACACT 1260
Db |||||
QY 2536 GCAACTCAAGTTGAACATCATCCGACAGACAGACAGACAGACAGACAGACAGACAGACACT 2595
Db |||||
QY 1261 CCTCTGCGACGACCGGACCGGACCCCAAGCAGAGAACCAACACAGCAGCAGCAGCAGCAGC 1320
Db |||||
QY 2596 CCTCTGCGACGACCGGACCGGACCCCAAGCAGAGAACCAACACAGCAGCAGCAGCAGCAGC 2655
Db |||||
QY 1321 ACTGACTTCTTGGAACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGCAGCAGCAGCAGC 1380
Db |||||
QY 2656 ACTGACTTCTTGGAACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGCAGCAGCAGCAGC 2715
Db |||||
QY 1381 AACACAACTCATCACCAGATACCGGAGAGAGTCCGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 1440
Db |||||
QY 2716 AACACAACTCATCACCAGATACCGGAGAGAGTCCGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 2775
Db |||||
QY 1441 TTAATTACCAATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 1500
Db |||||
QY 2776 TTAATTACCAATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 2835
Db |||||
QY 1501 CGA 1503
Db |||
QY 2836 AGA 2838
Db |||
```

RESULT 13

```
US-10-491-121-25
; Sequence 25, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
```

```
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 10783
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct PAD/CMV-OP(dTM) (Z-CITE-S)
US-10-491-121-25
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 10783;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATTGCAAGTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 60
Db |||||
QY 1324 ATGGCGGTTACAGGAATATTGCAAGTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 1383
Db |||||
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
Db |||||
QY 1384 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 1443
Db |||||
QY 121 AGCATTACAGGTTAGTGTGCGACAACTAGTTTGTGCGTGAACAACTGTGATCCACA 180
Db |||||
QY 1444 AGCATTACAGGTTAGTGTGCGACAACTAGTTTGTGCGTGAACAACTGTGATCCACA 1503
Db |||||
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db |||||
QY 1504 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 1563
Db |||||
QY 241 TCTGCNACTAAAGATGGGGCTTTCAGGTCCGGTCCCAACCAAGGTGGTCAATATGAA 300
Db |||||
QY 1564 TCTGCNACTAAAGATGGGGCTTTCAGGTCCGGTCCCAACCAAGGTGGTCAATATGAA 1623
Db |||||
QY 301 GCTGTGAATCGGCTGAAAACCTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGAG 360
Db |||||
QY 1624 GCTGTGAATCGGCTGAAAACCTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGAG 1683
Db |||||
QY 361 TGTCTACAGCAGCGCCAGCAGGAGTTGGGGCTTCCCGGTGCGGTATGTGCAAAA 420
Db |||||
QY 1684 TGTCTACAGCAGCGCCAGCAGGAGTTGGGGCTTCCCGGTGCGGTATGTGCAAAA 1743
Db |||||
QY 421 GTATCAGGAACGGACCGTGTGCGGAGACTTTCCTTCCATTAAGAGGGTGTCTTCTTC 480
Db |||||
QY 1744 GTATCAGGAACGGACCGTGTGCGGAGACTTTCCTTCCATTAAGAGGGTGTCTTCTTC 1803
Db |||||
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAACGACTTTCGCTGAAGGTGTC 540
Db |||||
QY 1804 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAACGACTTTCGCTGAAGGTGTC 1863
Db |||||
QY 541 GTTGCATTTCTGATATGTCGCCCAAGCTTAAGAGGACTTCTTTCAGCTCAACCCCTTGAGA 600
Db |||||
QY 1864 GTTGCATTTCTGATATGTCGCCCAAGCTTAAGAGGACTTCTTTCAGCTCAACCCCTTGAGA 1923
Db |||||
QY 601 GAGCGGTCAATGCAACGGAGAACCGGTCTAGTGGCTACTATTCTTACCAAAATTAGATAT 660
Db |||||
QY 1924 GAGCGGTCAATGCAACGGAGAACCGGTCTAGTGGCTACTATTCTTACCAAAATTAGATAT 1983
Db |||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGACTACTTGTTCGAGGTTTGACAATTGACC 720
Db |||||
QY 1984 CAGGCTACCGGTTTGGAAACCAATGAGACAGACTACTTGTTCGAGGTTTGACAATTGACC 2043
Db |||||
```

```
QY 721 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db |||||
QY 2044 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2103
Db |||||
QY 781 TATACAGTGGGAAAGAGGCAATACCAACGCGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
Db |||||
QY 2104 TATACAGTGGGAAAGAGGCAATACCAACGCGGAAACTAAATTTGGAAGGTCAACCCCGAA 2163
Db |||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAAACTTAAATAAACTCTACTAGAAAA 900
Db |||||
QY 2164 ATTGATACAACTCGGGAGTGGGCTTCTCGGAAACTTAAATAAACTCTACTAGAAAA 2223
Db |||||
QY 901 ATTCCAGTGAAGAGTTGTTTTCACAGTTGATCAACGCGGACCAAAACATCAGTGGT 960
Db |||||
QY 2224 ATTCCAGTGAAGAGTTGTTTTCACAGTTGATCAACGCGGACCAAAACATCAGTGGT 2283
Db |||||
QY 961 CAGAGTCCGGCGGAACTTCTCCGACCCAGGACCAACACAACTGAGACACACAA 1020
Db |||||
QY 2284 CAGAGTCCGGCGGAACTTCTCCGACCCAGGACCAACACAACTGAGACACACAA 2343
Db |||||
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db |||||
QY 2344 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 2403
Db |||||
QY 1081 GCAGTGTGCAATCTAACAAACCTTGCACAAATCTCCACAGTCCCAATCCCTCAACAC 1140
Db |||||
QY 2404 GCAGTGTGCAATCTAACAAACCTTGCACAAATCTCCACAGTCCCAATCCCTCAACAC 2463
Db |||||
QY 1141 AAACAGTGTCCGACAAACAGACCCNATAATACACCGCTGTATTAACCTTGACATCTCTGAG 1200
Db |||||
QY 2464 AAACAGTGTCCGACAAACAGACCCNATAATACACCGCTGTATTAACCTTGACATCTCTGAG 2523
Db |||||
QY 1201 GCAACTCAAGTTTGAACAACATCACCGCAACAGACAGACAGACAGACAGCTCCGACACT 1260
Db |||||
QY 2524 GCAACTCAAGTTTGAACAACATCACCGCAACAGACAGACAGACAGACAGCTCCGACACT 2583
Db |||||
QY 1261 CCCTCTGCCAGCACCGCAGCCGAGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 1320
Db |||||
QY 2584 CCCTCTGCCAGCACCGCAGCCGAGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 2643
Db |||||
QY 1321 ACTGACTTCTGGACCCCGCAGCACCAACAGTCCCAAAACACACGCGAGACCCGCTGGC 1380
Db |||||
QY 2644 ACTGACTTCTGGACCCCGCAGCACCAACAGTCCCAAAACACACGCGAGACCCGCTGGC 2703
Db |||||
QY 1381 AACAAACACACTCATCACCAAGATACCGAGAGAGAGTCCAGACAGCGGGAAGCTAGGC 1440
Db |||||
QY 2704 AACAAACACACTCATCACCAAGATACCGAGAGAGAGTCCAGACAGCGGGAAGCTAGGC 2763
Db |||||
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCGAGGACTGATCACAGCGGGGAGAGAACTCGA 1500
Db |||||
QY 2764 TTAATTACCAATACTATTGCTGGAGTCCGAGGACTGATCACAGCGGGGAGAGAACTCGA 2823
Db |||||
QY 1501 CGA 1503
Db 2824 AGA 2826
```

```
RESULT 14
US-09-337-946A-1
; Sequence 1, Application US/09337946A
; Patent No. US20020164582A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hart, Mary Katherine
; APPLICANT: Wilson, Julie A.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Ebola Virus Proteins Expressed from Venezuelan Equine Encephali
; TITLE OF INVENTION: Virus Replicons
; FILE REFERENCE: Army 144
; CURRENT APPLICATION NUMBER: US/09/337,946A
```

```
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 60/091,403
; FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: IBM compatible, Word 97, Windows 95
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola Zaire
US-09-337-946A-1
```

```
Query Match 66.6%; Score 1499.8; DB 9; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60
Db |||||
QY 59 ATGGCGGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 118
Db |||||
QY 61 CTTTGGGTAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTTGGAGTCAATCACAAT 120
Db |||||
QY 119 CTTTGGGTAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTTGGAGTCAATCACAAT 178
Db |||||
QY 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTCTGTCGACAAACTGTCTATCCACA 180
Db |||||
QY 179 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTCTGTCGACAAACTGTCTATCCACA 238
Db |||||
QY 181 AATCAATTTAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db |||||
QY 239 AATCAATTTAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
Db |||||
QY 241 TCTGCAACTAAAGAGTGGGGCTTTCAGGTCGCGTGTCCCAACCAAGGTGGTCAATTTATGAA 300
Db |||||
QY 299 TCTGCAACTAAAGAGTGGGGCTTTCAGGTCGCGTGTCCCAACCAAGGTGGTCAATTTATGAA 358
Db |||||
QY 301 GCTGTGTAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db |||||
QY 359 GCTGTGTAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 418
Db |||||
QY 361 TGCTTACACAGCAGCGCCAGACGGGATTCGGGCTTCCCGCGTGCCTGATGTGACACAA 420
Db |||||
QY 419 TGTCTACACAGCAGCGCCAGACGGGATTCGGGCTTCCCGCGTGCCTGATGTGACACAA 478
Db |||||
QY 421 GTATCAGGAAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db |||||
QY 479 GTATCAGGAAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 538
Db |||||
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGACGACTTTCGCTGAGGGTCTC 540
Db |||||
QY 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGACGACTTTCGCTGAGGGTCTC 598
Db |||||
QY 541 GTTGCATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db |||||
QY 599 GTTGCATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 658
Db |||||
QY 601 GAGCCGGTCAATGCAACCGGAGGACCCGCTAGTGCTACTATTCTACCAAAATAGATAT 660
Db |||||
QY 659 GAGCCGGTCAATGCAACCGGAGGACCCGCTAGTGCTACTATTCTACCAAAATAGATAT 718
Db |||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db |||||
QY 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 778
Db |||||
QY 721 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db |||||
QY 779 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 838
Db |||||
QY 781 TATACAAGTGGGAAAGAGGCAATACACCGGAAACTTAATTTGGAGGTCAACCCCGAA 840
Db |||||
QY 839 TATACAAGTGGGAAAGAGGCAATACACCGGAAACTTAATTTGGAGGTCAACCCCGAA 898
Db |||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAAACTTAAATAAACTCTACTAGAAAA 900
Db |||||
```

Db 899 ATTGATACAAATCGGGAGTGGCCCTTCTGGGAACTAAAAAAACCTCACTAGAAA 958
Qy 901 ATTCCAGTGAAGATTGTCTTTTCAAGTTGATCAAAAGGAGCCAAAAAATCATAGTGGT 960
Db 959 ATTCCAGTGAAGATTGTCTTTTCAAGTTGATCAAAAGGAGCCAAAAAATCATAGTGGT 1018
Qy 961 CAGAGTCGGCGGCAACTTCTCCGACCCAGGAGCCAAACAACTGAAGACCAAAA 1020
Db 1019 CAGAGTCGGCGGCAACTTCTCCGACCCAGGAGCCAAACAACTGAAGACCAAAA 1078
Qy 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 1079 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1138
Qy 1081 GCAGTGTGCTATTAACAAACCTTCCGCAATCTCCAGAGTCCCAATCCCTCAACAAC 1140
Db 1139 GCAGTGTGCTATTAACAAACCTTCCGCAATCTCCAGAGTCCCAATCCCTCAACAAC 1198
Qy 1141 AAACAGGTCCGGACAACAGCACCCATAATACACCGGTGTATAACTTGACATCTCTGAG 1200
Db 1199 AAACAGGTCCGGACAACAGCACCCATAATACACCGGTGTATAACTTGACATCTCTGAG 1258
Qy 1201 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGACAGACAGACAGCCCTCCGACACT 1260
Db 1259 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGACAGACAGACAGCCCTCCGACACT 1318
Qy 1261 CCCTCTGCAACAGCCGAGCGGAGCCCAAAAGCAGAGAACCAACACAGGAGCAAGAGC 1320
Db 1319 CCCTCTGCAACAGCCGAGCGGAGCCCAAAAGCAGAGAACCAACACAGGAGCAAGAGC 1378
Qy 1321 ACTGACTCTCTGGACCCCGCACCAACAGTCCCAAGTCCCAACCCAGCGAGACCGCTGGC 1380
Db 1379 ACTGACTCTCTGGACCCCGCACCAACAGTCCCAAGTCCCAACCCAGCGAGACCGCTGGC 1438
Qy 1381 AACAAACAACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1440
Db 1439 AACAAACAACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1498
Qy 1441 TTAATTAACAATATATTGCTGGAGTCGAGAGTGCAGAGTGCAGAGCGGAGAGAACTCGA 1500
Db 1499 TTAATTAACAATATATTGCTGGAGTCGAGAGTGCAGAGTGCAGAGCGGAGAGAACTCGA 1558
Qy 1501 CGA 1503
Db 1559 AGA 1561

RESULT 15

US-10-384-976-1

; Sequence 1, Application US/10384976
; Publication NO. US20030224015A1
; GENERAL INFORMATION:
; APPLICANT: HART, MARY KATHERINE
; APPLICANT: WILSON, JULIE A.
; APPLICANT: PUSKO, PETER
; APPLICANT: SMITH, JONATHAN F.
; APPLICANT: SCHMALJOHN, ALAN L.
; TITLE OF INVENTION: EBOLA PEPTIDES AND IMMUNOGENIC COMPOSITIONS CONTAINING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: ARMY 144A
; CURRENT APPLICATION NUMBER: US/10/384,976
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 09/337,946
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola zaire
US-10-384-976-1

Query Match 66.6%; Score 1499.8; DB 17; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCGGTTCAGGAAATATTGACGTTACCTCGTGCATCGATTCAAGAGGACATCATTTCTTT 60
Db 59 ATGGCGGTTCAGGAAATATTGACGTTACCTCGTGCATCGATTCAAGAGGACATCATTTCTTT 118
Qy 61 CTTTGGGTAAATATCTCTTTTCCAAAGAAACATTTTCCATCCCATCTGGAGTCATCCACAAT 120
Db 119 CTTTGGGTAAATATCTCTTTTCCAAAGAAACATTTTCCATCCCATCTGGAGTCATCCACAAT 178
Qy 121 AGCATTATACAGGTAGTGTGCGACAACTAGTTTGTGCGTGAACAACTGTCTATCCACA 180
Db 179 AGCATTATACAGGTAGTGTGCGACAACTAGTTTGTGCGTGAACAACTGTCTATCCACA 238
Qy 181 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 239 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
Qy 241 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTGTCCACCCAAAGGTGGTCAATATATGAA 300
Db 299 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTGTCCACCCAAAGGTGGTCAATATATGAA 358
Qy 301 GCTGTTGAATGGGCTGAAAACCTGCTACAACTTTGAAATCAAAAAACCTGACGGAGTGAG 360
Db 359 GCTGTTGAATGGGCTGAAAACCTGCTACAACTTTGAAATCAAAAAACCTGACGGAGTGAG 418
Qy 361 TGTCTACAGCAGCGCCAGCGGATTCGGGGCTTCCCGCGTCCCGGTATGTGCACAAA 420
Db 419 TGTCTACAGCAGCGCCAGCGGATTCGGGGCTTCCCGCGTCCCGGTATGTGCACAAA 478
Qy 421 GTATCAGGAAACCGGACCGTGTGCCGAGACTTTCCCTTCCATAAAGAGGGTCTTTCTTC 480
Db 479 GTATCAGGAAACCGGACCGTGTGCCGAGACTTTCCCTTCCATAAAGAGGGTCTTTCTTC 538
Qy 481 CTGTATGATCGACTTGTCTCCACAGTTATCTACCGAGGAGACGACTTTCCGCTGAAGGTGTC 540
Db 539 CTGTATGATCGACTTGTCTCCACAGTTATCTACCGAGGAGACGACTTTCCGCTGAAGGTGTC 598
Qy 541 GTTCATTTCTGATACCTCCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 599 GTTCATTTCTGATACCTCCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 658
Qy 601 GAGCGGTCAATGCAACGGAGACCCGCTAGTGGCTACTATTCTACCAATTAGATAT 660
Db 659 GAGCGGTCAATGCAACGGAGACCCGCTAGTGGCTACTATTCTACCAATTAGATAT 718
Qy 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTTGTTCCGAGGTTGACAAATTTGACC 720
Db 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTTGTTCCGAGGTTGACAAATTTGACC 778
Qy 721 TACGTCCAACTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGAATGAGACATA 780
Db 779 TACGTCCAACTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGAATGAGACATA 838
Qy 781 TATCAAGTGGGAAAGAGCAATACCAACGGGAAACCTAAATTTGGAGGTCACACCCCGAA 840
Db 839 TATCAAGTGGGAAAGAGCAATACCAACGGGAAACCTAAATTTGGAGGTCACACCCCGAA 898
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAACCCTCCTAGAAAAA 900
Db 899 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAACCCTCCTAGAAAAA 958
Qy 901 ATTGCGAGTGAAGATTGTCTTTTCCAGTGTCTATCAACGGGAGCCAAAAAATCATAGTGGT 960
Db 959 ATTGCGAGTGAAGATTGTCTTTTCCAGTGTCTATCAAAACGGGAGCCAAAAAATCATAGTGGT 1018
Qy 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCCAAACAACTGAAGAGACCAAAA 1020
Db 1019 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCCAAACAACTGAAGAGACCAAAA 1078

Qy	1021	ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1080
Db	1079	ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1138
Qy	1081	GCAGTGTGCACTTAACAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACCC	1140
Db	1139	GCAGTGTGCACTTAACAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACCC	1198
Qy	1141	AAACGAGGTCCGGACAAACAGCACCCCATATACACCCGTGTATATAACTTGACATCTCTGAG	1200
Db	1199	AAACGAGGTCCGGACAAACAGCACCCCATATACACCCGTGTATATAACTTGACATCTCTGAG	1258
Qy	1201	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGACAGCCTCCGACACT	1260
Db	1259	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGCCTCCGACACT	1318
Qy	1261	CCCTTGCCACGACCGCAGCCGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	1320
Db	1319	CCCTTGCCACGACCGCAGCCGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	1378
Qy	1321	ACTGACTTCTGACCCCGCCACCAACAAGTCCCAAAACACAGCGAGAGACCGCTGGC	1380
Db	1379	ACTGACTTCTGACCCCGCCACCAACAAGTCCCAAAACACAGCGAGAGACCGCTGGC	1438
Qy	1381	AACAACAACACTCATCAACAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC	1440
Db	1439	AACAACAACACTCATCAACAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC	1498
Qy	1441	TTAATTACCAATCTATTCTGGAGTCGAGGACTGATCACAGGCGGGAGAGAACTCGA	1500
Db	1499	TTAATTACCAATCTATTCTGGAGTCGAGGACTGATCACAGGCGGGAGAGAACTCGA	1558
Qy	1501	CGA.1503	
Db	1559	AGA.1561	

Search completed: August 1, 2005, 20:40:55
Job time : 1421.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 7912.14 Seconds

(without alignment)
11274.581 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctct.....ttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1841	100.0	1841	6	AX717715	AX717715 Sequence
2	1308	71.0	2046	6	AX717717	AX717717 Sequence
3	1303.4	70.8	11460	6	AR279677	AR279677 Sequence
4	1303.4	70.8	19104	14	MVREPCYC	Z12132 Marburg vir
5	1301.8	70.7	19112	14	AY430366	AY430366 Lake Vict
6	1301.8	70.7	19113	14	AY430365	AY430365 Lake Vict
7	1274	69.2	2051	6	AX717721	AX717721 Sequence
8	1151.4	62.5	2046	14	AF005735	AF005735 Marburg v
9	1148.2	62.4	2948	14	MAVSPAA	X68493 Marburg Vir
10	1148.2	62.4	19112	14	MVIRPR	Z29337 Marburg vir
11	1098.6	59.7	2046	14	AF005733	AF005733 Marburg v
12	1098.6	59.7	19151	14	AY358025	AY358025 Marburg v
13	780	42.4	2046	6	AX717719	AX717719 Sequence
14	780	42.4	2046	6	AX717725	AX717725 Sequence
15	777	42.2	2046	14	AF005734	AF005734 Marburg v
16	540	29.3	2039	6	AX717723	AX717723 Sequence
17	527.6	28.7	2406	14	EVU31033	U31033 Zaire Ebola
18	527.6	28.7	2408	14	EVU23187	U23187 Zaire Ebola
19	527.6	28.7	11742	14	EBORNA	L11365 Zaire Ebola

20	527.6	28.7	18959	14	AF086833	AF086833 Zaire Ebo
21	527.6	28.7	18959	14	AF272001	AF272001 Zaire Ebo
22	526	28.6	2298	6	AR404895	AR404895 Sequence
23	526	28.6	2298	6	AX092102	AX092102 Sequence
24	526	28.6	18959	14	AY142960	AY142960 Zaire Ebo
25	526	28.6	18960	14	AF499101	AF499101 Zaire Ebo
26	524.4	28.5	2408	14	EVU81161	EVU81161 Zaire Ebola
27	520.8	28.3	2408	14	EVU28077	EVU28077 Zaire Ebola
28	520.8	28.3	18961	14	AY354458	AY354458 Zaire ebo
29	516.4	28.0	2446	14	EVU77384	EVU77384 Zaire Ebola
30	514.8	28.0	2173	14	AY058898	AY058898 Zaire Ebo
31	511.6	27.8	2173	14	AY526098	AY526098 Zaire ebo
32	511.6	27.8	2173	14	AY526099	AY526099 Zaire ebo
33	511.6	27.8	2173	14	AY526100	AY526100 Zaire ebo
34	511.6	27.8	2173	14	AY526101	AY526101 Zaire ebo
35	511.6	27.8	2173	14	AY526102	AY526102 Zaire ebo
36	511.6	27.8	2173	14	AY526104	AY526104 Zaire ebo
37	510	27.7	1041	14	AY526103	AY526103 Zaire ebo
38	510	27.7	2173	14	AY526105	AY526105 Zaire ebo
39	324.2	17.6	2408	14	EVU28006	U28006 Cote d'Ivoi
40	286.6	15.6	2033	14	EVU23416	U23416 Reston Ebo1
41	286.6	15.6	2033	14	EVU23417	U23417 Reston Ebo1
42	286.6	15.6	2233	14	AF034645	AF034645 Ebola vir
43	286.6	15.6	2360	14	EVU23152	U23152 Reston Ebo1
44	286.6	15.6	18891	14	AF522874	AF522874 Reston Eb
45	283.4	15.4	18890	14	AB050936	AB050936 Reston Eb

ALIGNMENTS

RESULT 1	AX717715	AX717715	1841 bp	DNA	linear	PAT 15-APR-2003
LOCUS	Sequence 3 from Patent WO02079239.					
DEFINITION	AX717715					
ACCESSION	AX717715.1	GI:29890726				
VERSION						
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.					
TITLE	Chimeric filovirus glycoprotein					
JOURNAL	Patent: WO 02079239-A 3 10-OCT-2002;					
FEATURES	U.S. Army Medical Research Institute of Infectious Diseases (US)					
source	Location/Qualifiers					
	1..1841					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="Chimeric molecule between Marburg virus strain					
	Musoke Glycoprotein 1 and Ebola virus Zaire Mayinga strain					
	Glycoprotein 2"					

ORIGIN	Query Match	100.0%;	Score 1841;	DB 6;	Length 1841;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1841;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATCAAGGGACAAAAATCTCCCC	60		
Db	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATCAAGGGACAAAAATCTCCCC	60		
Qy	61	ATTTTAGAGTAGCTAGTAATAACCCCAAAATGGGATTCGGTATGTCCTCGGAAT	120		
Db	61	ATTTTAGAGTAGCTAGTAATAACCCCAAAATGGGATTCGGTATGTCCTCGGAAT	120		
Qy	121	CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180		
Db	121	CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180		
Qy	181	GATTCCCTTTGGAGGGCATCCAGGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT	240		

181 GATTCCTCTTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAA 240
241 GTTGAGTACACAGAGGGGAGGAGCCAAAACATGCTACAAATATAAGTGTACAGGATCCC 300
241 GTTGAGTACACAGAGGGGAGGAGCCAAAACATGCTACAAATATAAGTGTACAGGATCCC 300
301 TCTGAAAATCCTTGTTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
301 TCTGAAAATCCTTGTTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
361 ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
361 ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACATCCGTGACTATCCGAAATGCAAA 480
421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACATCCGTGACTATCCGAAATGCAAA 480
481 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAAGACAGTGCACAAAATGATTTCTCG 540
481 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAAGACAGTGCACAAAATGATTTCTCG 540
541 CGGCAAGGACAAAGGTACCGTCAATATGAATCTGACTTCTAATAAATATTTGGACAAGT 600
541 CGGCAAGGACAAAGGTACCGTCAATATGAATCTGACTTCTAATAAATATTTGGACAAGT 600
601 ACTACGGAAACCAACGAATGACACTGATGTTTGGGGCTCTTCAAGAAATACAAATCT 660
601 ACTACGGAAACCAACGAATGACACTGATGTTTGGGGCTCTTCAAGAAATACAAATCT 660
661 ACAAGAACCAACATGCTCGTCCAAATACCTCCACACATGCTCCACAGCCCGTCCG 720
661 ACAAGAACCAACATGCTCGTCCAAATACCTCCACACATGCTCCACAGCCCGTCCG 720
721 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 780
721 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 780
781 AGTGATGATGAGGACCTCCGACATCCGCTCAGGTCGGGAGACGAAACCCCAACACA 840
781 AGTGATGATGAGGACCTCCGACATCCGCTCAGGTCGGGAGACGAAACCCCAACACA 840
841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCAACAAATGCTCCACCTCCCTCACCA 900
841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCAACAAATGCTCCACCTCCCTCACCA 900
901 CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAACCATTTCCCAAGATGCTGTGACT 960
901 CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAACCATTTCCCAAGATGCTGTGACT 960
961 GAACTAGACAAAATAACACACTGACAAACCGTCCATGCCCCCTCATAACTACCACA 1020
961 GAACTAGACAAAATAACACACTGACAAACCGTCCATGCCCCCTCATAACTACCACA 1020
1021 ATCTCTACTAAACAACCTCCAAACAACACTTTCAGCACTCTCTGCAACCATTAACAAC 1080
1021 ATCTCTACTAAACAACCTCCAAACAACACTTTCAGCACTCTCTGCAACCATTAACAAC 1080
1081 ACCACCAATGACAAACACAGAGCACAATCACTGAAAATGAGCAACACAGTCCCTCCG 1140
1081 ACCACCAATGACAAACACAGAGCACAATCACTGAAAATGAGCAACACAGTCCCTCCG 1140
1141 ATACCAACCTGCTCCAGGGGAATCCACACAGCAAGAGGACGAGCAAGGACCAAAA 1200
1141 ATACCAACCTGCTCCAGGGGAATCCACACAGCAAGAGGACGAGCAAGGACCAAAA 1200
1201 GGCCCCGACCAACAGGCAACCAAGATGAGCATTTTCAACAGTCTCTCCCTCCAC 1260
1201 GGCCCCGACCAACAGGCAACCAAGATGAGCATTTTCAACAGTCTCTCCCTCCAC 1260
1261 CCAGCTCGACTGCAACAACATCTGTATATTTTCAAGAAAGCGATTCGGCAATTTGTCAAT 1320

Db 1261 CCCAGCTCGATGCGACAACATCTTGTATATTTTCAAGAAAGCGATCGGCAATTTGTCAAT 1320
QY 1321 GCTCAACCCAAATGCAACCCCTAAATTTTACATTTACTGGACTACTCAGGATGAAGTGTCTGCA 1380
Db 1321 GCTCAACCCAAATGCAACCCCTAAATTTTACATTTACTGGACTACTCAGGATGAAGTGTCTGCA 1380
QY 1381 ATCGGACTGGCTGGATACCATATTTTCGGGGCAGCAGCGAGGGAATTTACATAGAGGG 1440
Db 1381 ATCGGACTGGCTGGATACCATATTTTCGGGGCAGCAGCGAGGGAATTTTACATAGAGGG 1440
QY 1441 CTAATGCACAATCAAGATGGTTTAAATCTGTGGTTTGAAGAGTGTGCGCAACGAGAGGACT 1500
Db 1441 CTAATGCACAATCAAGATGGTTTAAATCTGTGGTTTGAAGAGTGTGCGCAACGAGAGGACT 1500
QY 1501 CAAGCTCTTCAACTGTTTCTCTGAGAGCCAACTGAGCTACGCACTTTTCAATCCTCAAC 1560
Db 1501 CAAGCTCTTCAACTGTTTCTCTGAGAGCCAACTGAGCTACGCACTTTTCAATCCTCAAC 1560
QY 1561 CGTAAGGCAATTTGATTTCTTCTGAGCGATGGGGCGGCACATGCCACATTTCTGGAGCG 1620
Db 1561 CGTAAGGCAATTTGATTTCTTCTGAGCGATGGGGCGGCACATGCCACATTTCTGGAGCG 1620
QY 1621 GACTGCTCTATCGAACCATGATTTGGACCAAGACATTAACAGACAAAATTTGATCAGATT 1680
Db 1621 GACTGCTCTATCGAACCATGATTTGGACCAAGACATTAACAGACAAAATTTGATCAGATT 1680
QY 1681 ATTCATGATTTTGTGATAAAACCTTCCGACCAAGGGGACAAATGCAAAATTTGGTGACA 1740
Db 1681 ATTCATGATTTTGTGATAAAACCTTCCGACCAAGGGGACAAATGCAAAATTTGGTGACA 1740
QY 1741 GGATGAGACAAATGATACCGGACAGGATTTGAGTTTACAGGGCTTATATTCGAGTTATC 1800
Db 1741 GGATGAGACAAATGATACCGGACAGGATTTGAGTTTACAGGGCTTATATTCGAGTTATC 1800
QY 1801 GCTTTATTTCTGATATGCAAAATTTGCTCTTTAGTTGAATTC 1841
Db 1801 GCTTTATTTCTGATATGCAAAATTTGCTCTTTAGTTGAATTC 1841

RESULT 2
AX717717
LOCUS AX717717 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS .
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1. .2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Musoke glycoprotein 1 and Marburg virus strain Raven
Glycoprotein 2"

ORIGIN
Query Match 71.0%; Score 1308; DB 6; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAATCTCCCC 60
QY 61 ATTTTAGATAGTAGTAGTAATAATCAACCCCAAAATGTGGATTGGTCTCCGGAAT 120

[illegible]

Db	6420	ACTGAAGGGAACATAGCAGCTATGATTGTCATTAAGACAGCTGCACAAATGATTTTCTCG	6479		
QY	541	CGGCAAGGACAAAGGGTACCGTCAATGAATCTGACTTCTACTAATAAATATTGGACAAGT	600		
Db	6480	CGGCAAGGACAAAGGGTACCGTCAATGAATCTGACTTCTACTAATAAATATTGGACAAGT	6539		
QY	601	AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATCT	660		
Db	6540	AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATCT	6599		
QY	661	ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACCTTGCCCAACAGCCCGTCCG	720		
Db	6600	ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACCTTGCCCAACAGCCCGTCCG	6659		
QY	721	GAGATCAAACTCACAAGCACCCCAACTGATGCACCAAACTCAATACCAAGCACCCCAAGC	780		
Db	6660	GAGATCAAACTCACAAGCACCCCAACTGATGCACCAAACTCAATACCAAGCACCCCAAGC	6719		
QY	781	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACA	840		
Db	6720	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACA	6779		
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACATGCCACCACTCCCTCACCA	900		
Db	6780	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACATGCCACCACTCCCTCACCA	6839		
QY	901	CAACCAAGCACGCCACAGCAAGGAGGAAACAACAACATTTCCCAAGATGCTGTGACT	960		
Db	6840	CAACCAAGCACGCCACAGCAAGGAGGAAACAACAACATTTCCCAAGATGCTGTGACT	6899		
QY	961	GAACCTAGACAAAATAACAACTGCACAAACCGTCCATGCCCTCATTAACACTACCAACA	1020		
Db	6900	GAACCTAGACAAAATAACAACTGCACAAACCGTCCATGCCCTCATTAACACTACCAACA	6959		
QY	1021	ATCTCTACTAACAACACCTCCAAACACTTTCAGCACTCTCTCTGACCACTTACAAAAC	1080		
Db	6960	ATCTCTACTAACAACACCTCCAAACACTTTCAGCACTCTCTCTGACCACTTACAAAAC	7019		
QY	1081	ACCACCAATGACAAACACAGACGACAACTCACTGAAATGAGCAAAACAGTGCCTCCCTCG	1140		
Db	7020	ACCACCAATGACAAACACAGACGACAACTCACTGAAATGAGCAAAACAGTGCCTCCCTCG	7079		
QY	1141	ATAACAACCTGCTCCAAACGGGAATCCCAACACAGCAAAAGAGCACAGCAGCAAAAA	1200		
Db	7080	ATAACAACCTGCTCCAAACGGGAATCCCAACACAGCAAAAGAGCACAGCAGCAAAAA	7139		
QY	1201	GGCCCCGCGCAACAGGGCACAAACAGCAAAATGAGCATTTTCAACAGTCTCTCCCTCACC	1260		
Db	7140	GGCCCCGCGCAACAGGGCACAAACAGCAAAATGAGCATTTTCAACAGTCTCTCCCTCACC	7199		
QY	1261	CCAGCTCGACTGCACAACTCTGTATATTTTCAGAGAAAGCGA	1305		
Db	7200	CCAGCTCGACTGCACAACTCTGTATATTTTCAGAGAAAGCGA	7244		
RESULT 5	AY430366	19112 bp	RNA	linear	VRL 06-FEB-2004
LOCUS	AY430366	Lake Victoria marburgvirus strain pp4 guinea pig nonlethal variant, complete genome.			
DEFINITION	AY430366	AY430366.1 GI:40388379			
ACCESSION	AY430366	Lake Victoria marburgvirus			
VERSION	AY430366	Lake Victoria marburgvirus			
KEYWORDS	AY430366.1	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.			
SOURCE	AY430366.1	1 (bases 1 to 19112)			
ORGANISM	AY430366.1	Hevey, M., Negley, D., Geisbert, J., Jahrling, P. and Schmaljohn, A.			
REFERENCE	AY430366.1	Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants			
AUTHORS	AY430366.1	Virology 239 (1), 206-216 (1997)			
TITLE	AY430366.1	98087840			
JOURNAL	AY430366.1				
MEDLINE	AY430366.1				

9426460	2	(bases 1 to 19112)	
REFERENCE	Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H., Smith, K.L. and McCready, P.M.		
AUTHORS	Direct Submission		
TITLE	Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA		
JOURNAL	3 (bases 1 to 19112)		
REFERENCE	Hevey, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Lofes, L., Ichou, M. and Schmaljohn, A.		
AUTHORS	Direct Submission		
TITLE	Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort Detrick, MD 21702, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1. 19112		
source	/organism="Lake Victoria marburgvirus"		
	/mol_type="genomic RNA"		
	/strain="pp4 guinea pig nonlethal variant"		
	/isolate="Musoke Kenya, 1980"		
	/db_xref="taxon:11269"		
gene	104..2191		
	/gene="NP"		
CDS	104..2191		
	/gene="NP"		
	/note="nucleoprotein"		
	/codon_start=1		
	/product="NP protein"		
	/protein_id="AAR85453.1"		
	/db_xref="GI:40388380"		
	/translation="MDLSLLELGTGTAPHRNKKVILFTNHQVSCNQIIDAIS GIDGDLLEGGLTLCVEHYNSDKFNTSPYAKYRDAGYEFVDYIKNADRELDV IPNEPHYSPIILALKLTLESTESQRIGLFSLFLPKLVGDRASIEKALRQVTV HOEGIVTVYPNHMLTGMKVIIGILRSSFLKFLVLIHQGNLVTGHDAYSIIISNV GQTFESGLIVKTVLETLQKDSGLVTLPLVRTSKVNEVASPKQALNSLARHGEYA PFAVNLGSLNLEHGLYFOLSAIAGLVATAHGSTLAGVNVGBOYQQLREAHDAEV KLQRHHEQEIQAIAEDDEKILQFHLQKTEITHSQTALVLSQKREKLARAEIE NNIVEDGQKFSQNSQSLFNDTPVEVTQVAPMNRPTALPPVDDKIBHESTEDS SSSSFDVLDNPFALLNEDDSDSVIMPGTTSREFGQIPEPRQSDLNNSQKQE DESTNPKQLFRLYQELPPVOEDDESVTTDSQESIDQSGDNQGVLDLPPPLYAQE KROPIOHPRANPODPFGSIGDVNGDILEPIRSPSSPSAPOEDTMRAYELSPFTN DEDNQNPQPVVTKGRTFLYFNLDLQTNPPESLITAVEYQNPVSAKELQADMPD MSFDRHRVAVNL"		
gene	2945..3934		
	/gene="VP35"		
CDS	2945..3934		
	/gene="VP35"		
	/note="minor component of RNA-dependent RNA-polymerase"		
	/codon_start=1		
	/product="VP35 protein"		
	/protein_id="AAR85454.1"		
	/db_xref="GI:40388381"		
	/translation="MWDSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRPKCTVG LQCSPLMSKATSDDDIWDQLIVKRTLADLLIPINQISDIQSTLSEVTVRHEIER QLHEITPVLMKGRTELEAISKGMSEMLAKIDHLVISTGRITAPAAAFDAYLNEHGVPPP QPAIFKDLGVAQAQCSKGTMYKNATDAADKMSKVLSESETFKSPKLSAKDLALLF THLPNNTPPHILAQVLSKIAKSGKGAFLDAFHQILSEGENAALTRLSRTFDAF LGVVPVIRVKNFQTPRCPCKSLRAVPNPNTIDKGWVCVYSSEQGETRALKI"		
gene	4568..5479		
	/gene="VP40"		
CDS	4568..5479		
	/gene="VP40"		
	/note="matrix protein"		
	/codon_start=1		
	/product="VP40 protein"		
	/protein_id="AAR85455.1"		
	/db_xref="GI:40388382"		
	/translation="MASSSNYNTYQVLYNPPFYADHGANQLIPADQLNSQOQITPNVY GDLLDDQFGKNVCHAFTELEAIIISAYNRTVKGVPAWPLFLGMSNFYPLAHFVAA TLTGSTVLTQPTNGQKFRVIRHPLRLMREGNQAFIQNVIPRNFSTNQK LYNTNLVTSVKQLPDDAWRPSKDKLIGNTHWPAVSIPNLPIIPLPTVTKQAYRQHK NPNNGPLLAISGILHQLRVEKVPKTSLSFRISLPADMFVSKGMMKKEGNSPVVYFQ APENFPFLGFMNRQVVLAYANPLSVA"		
gene	5941..7998		

QY	661	ACAAAGAACAAACATGTCTCGTCCAAATACTCTCCACCACTGCCACAGCCCGTCCG	720
Db	6601	ACAAAGAACAAACATGTCTCGTCCAAATACTCTCCACCACTGCCACAGCCCGTCCG	6660
QY	721	GAGATCAAACTCACAAGACCCCACTGATGCGCAGCAAACTCAATACCGAGCCCAAGC	780
Db	6661	GAGATCAAACTCACAAGACCCCACTGATGCGCAGCAAACTCAATACCGAGCCCAAGC	6720
QY	781	AGTGATGATGAGGACCTCCGCAACATCCGCTCAGGTCGGGAGAGAGAGAAACCCACACA	840
Db	6721	AGTGATGATGAGGACCTCCGCAACATCCGCTCAGGTCGGGAGAGAGAGAAACCCACACA	6780
QY	841	ACTTGTGATGCGGTCAACAGCAAGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA	900
Db	6781	ACTTGTGATGCGGTCAACAGCAAGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA	6840
QY	901	CAACCAAGCAGCCACAGCAAGGAGGAGAAACACACAAACCATTTCCCAAGATCTGTGACT	960
Db	6841	CAACCAAGCAGCCACAGCAAGGAGGAGAAACACACAAACCATTTCCCAAGATCTGTGACT	6900
QY	961	GAACTAGACAAAAATAACAACTGACAAACCGTCCATGCCCTCCCTCATAACTACCA	1020
Db	6901	GAACTAGACAAAAATAACAACTGACAAACCGTCCATGCCCTCCCTCATAACTACCA	6960
QY	1021	ATCTCTACTAACCAACCTCCAAACAACTTTCAGCACTCTCTCTGCACCACTTACAAAC	1080
Db	6961	ATCTCTACTAACCAACCTCCAAACAACTTTCAGCACTCTCTCTGCACCACTTACAAAC	7020
QY	1081	ACCACCAATGACACACAGAGCAACATCACTGAAATGACAAACCACTGCTCCCTCCG	1140
Db	7021	ACCACCAATGACACACAGAGCAACATCACTGAAATGACAAACCACTGCTCCCTCCG	7080
QY	1141	ATAACAAACCTGCTCCCAACGGAATCCACACAGCAAGCAGCAGCAAGCAAAA	1200
Db	7081	ATAACAAACCTGCTCCCAACGGAATCCACACAGCAAGCAGCAGCAAGCAAAA	7140
QY	1201	GGCCCCGACCAACGCGCACCACCAACAGCAAAATGAGCATTTCCAGCTCTCTCCCCACC	1260
Db	7141	GGCCCCGACCAACGCGCACCACCAACAGCAAAATGAGCATTTCCAGCTCTCTCCCCACC	7200
QY	1261	CCAGCTCAGCTGACACAACTTGTATATTTTCAGAGAAAGCGA	1305
Db	7201	CCAGCTCAGCTGACACAACTTGTATATTTTCAGAGAAAGCGA	7245
RESULT 6			
AY430365		19113 bp	RNA
LOCUS	AY430365	Lake Victoria marburgvirus strain pp3 guinea pig lethal variant, complete genome.	VR_L 06-FEB-2004
DEFINITION	AY430365		
ACCESSION	AY430365.1	GI:40388387	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

TITLE	Hevey, M. and Schmaljohn, A.
JOURNAL	Direct Submission
FEATURES	Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort Detrick, MD 21702, USA
source	Location/Qualifiers
	1..19113
	/organism="Lake Victoria marburgvirus"
	/mol_type="genomic RNA"
	/strain="pp3 guinea pig lethal variant"
	/isolate="Musoke Kenya, 1980"
	/db_xref="taxon:11269"
gene	104..2191
	/gene="NP"
CDS	104..2191
	/note="nucleoprotein"
	/codon_start=1
	/product="NP protein"
	/protein_id="AAR85460.1"
	/db_xref="GI:40388388"
	/translation="MDLHSLLELGTPTAPVRNKKVILFDTNHQVSIQNIIIDAINS GIDGLLEGLLTLCEVHYNSDKFNTSPYAKYLRDAGYEDVIKNADATRELDV IPNEPHYSPLILALKLTLESTESQGRIGLFLSCSLPLKLVVGDRAISIALROVTV HQEGIVTYPNHWTTHGMKVIKILASSPILKVLHOGVNLVTHGDAYDSILNSV GTRPSGLLIVKTVLEILQKTSVTLHPLVRISKVNEVASFKQALNSLARHGEYA PFARVNLGSLNLEHGLYPQLSAIALGVATAHGSTLAGVNVGEYQQLRRAAHDAEV KLQRHHEQEIQAIEDDEERKILEQFLQKTEITHSQTLAVLSQKREKLARAAEIE NNIVDDQCFKOSQNRVSFLNDPTPEVTVOARPMNRPTALPPVDDKI BHESTEDS SSSSVDLNDPPALLNEDEDTLDDSVMPCTTSREFOGIPEPQSDLNNSQKQOE DESTNPIKKFLRYQELPPVQEDDESEYTTDSQSIDQPGSDNEGVDLPPPLVAQE KRQDIQHPAANPQDPFGSDGVNDIDLEIRSPSPSPAPQEDTMRRAELSPDTFN DEDQNNPQRVVTGKRTFLYPNDLLQTNPPESLITLALVEEYQNPVSAKELQADWPD MSFDERRHVAMNL"
gene	2945..3934
	/gene="VP35"
CDS	2945..3934
	/note="minor component of RNA-dependent RNA-polymerase"
	/codon_start=1
	/product="VP35 protein"
	/protein_id="AAR85461.1"
	/db_xref="GI:40388389"
	/translation="MWDSSYMQQVSEGMLTKGVIPIDQVFGANPLKLYKRRKPKGTVG LQSPCLMSKATSTDDIIMDQLIVKRTLADLLIPINRQISDIQSTLSVTTVRHEIER QLEHITVLMKGRLEAISKGSEMELAKYDHLVISTGRTTAPAAAFDAYLNEHGVPPP QPAIFKDLGVAAQACSKGTVMKNATTDAAKMSKVLSEETFKPNLSAKDLALLF THLPQNTTPEHILAQVLSKIAYKSGSGAFILDAFHQILSEGENAQALTRLSTDFAF LCVVPVIRKVNQTVPRPCQKSLRAVPPNPTIDKGNVCVYSSESQGETRALKI"
gene	4568..5479
	/gene="VP40"
CDS	4568..5479
	/note="matrix protein"
	/codon_start=1
	/product="VP40 protein"
	/protein_id="AAR85462.1"
	/db_xref="GI:40388390"
	/translation="MASSSYNTVMQYLNPPPYADHGANOLIPADLSNQOIGTPNYV GDLNDDQFGKGVCHAFLEAIIIDISAYNETVKGVPAWLPLGIMSNFEYPLAHTVAA LYTQSYLTQTHNGKQFVRNRLGTGIPAPHLRLMREGNOAFTQNMVLPNFSNQF TNINVLNLSVKQLPDDAWRPSKNKLGINTWHPAVSHPNLPPIVLPTVKQAYKQHK NPNNGPLAIISGILHQRVEKVPKTSLSFRISLPADMFVKEGMKKRGKGENSPVYVQ APENFPLNGFNRRQVLAAYANPTLSAV"
gene	5941..7986
	/gene="GP"
CDS	5941..7986
	/note="envelope transmembrane glycoprotein"
	/codon_start=1
	/product="GP protein"
	/protein_id="AAR85463.1"
	/db_xref="GI:40388391"
	/translation="MKTTCFLISLILIQGTKNLFILEASNNQPNQVDSVCSGTQKLT"

EDVHLMGFTLSGQKVDASPLASKEWAFRTGVPKKNVRYTGERAKTCYNI SVTDPSG
KSLLLDPTNIRDYKCKTIIHQONPHAGIOLWGAFFLDRIASTMYRGKVF
TEGNTAAMI VNTYKHOMLFSKQGGIRHNNLTSTNKYTSNGQNTDNCFCGALQY
NSTKQTCAPSKIPPLTPARPEIKLTSTDTATKLTNTDPSDDSDIATSGSGGER
EPIHTTSADVTKGLSTMPPTSPQSPSTPOQGNNTNHSQDAVTELDNNTTAQPSMP
PHNTTTSIGNTNSKHNFSTLAPLONTNDNTQSTITENEOTSAPIITLTPPTGNPT
AKSTSSKGTATTA PNTTNEHTSPPTPSSTAQHLVFRKRSILWREGGPFLLDG
LNAIDPDPVPTNTKIIDESSSSGASABEDQHASPNISLISLSPNNINENTAYSGEN
ENDCAELRIMSVDQDDLAAGLSWIFFPGGLEGTAGLGNKQNDLKRRLANQT
AKSLSELPNTTEERTSLINRHAIDFLLTGWGTCVKLGPDCQIGIEDLSKRSIQEI
QKDKDEQEGTGMGLGQKMTSDMGVLTNLGILLILLIAVLIALSCICRIETFYIG"
8869. .9714
/gene="VP30"
8869. .9714
/gene="VP30"
/codon_start=1
/product="VP30 protein"
/protein_id="AAR85464.1"
/db_xref="GI:40388392"
/translation="MQQPRGRSRTNRHOVPTIYHETOLPSKPHYNTYHPRARMSST
RSSAESPTNHIIPRAPSPSTFNLSKPPPPKMDCKNMKIGLPCADPTCNRDHDLNLT
NRELLLMARMLNPTDKTFRSPQCGSPSLSKGLSKDKQKQDTKDLVLENLHLSY
LHRSIGKDTLSLAALSLTGAGIRKTRNSLTINTMTLHNNHENLPQDQNGVAKQTY
TGHLDKGQPEALWQGDWKRSLSLFVQALYVNNNIPCESSISVQASVDHFLIPQS
QCKGO"
10207. .10968
/gene="VP24"
10207. .10968
/gene="VP24"
/codon_start=1
/product="VP24 protein"
/protein_id="AAR85465.1"
/db_xref="GI:40388393"
/translation="MAELSTRYNLPANVTENSINDLNSTARWIKEPSVGOWTVKWN
GVFPHIENGTMTLHLKNSFVLPVPEQOQNTFLSHLPKPKSTIIEPFLAILLGLVAL
KQELQOQSLPGRSIVHMLSEWLLVETSAIHISPNLIGVLTSDMFKILMAGVNF
ENKMFLLVANDHKGPSLEIKLTGQOILITRVNMGFLVEVRRIDIEFCCEGTVLS
VYGLVAAVLAREHSEQMEKQGPLNLTQMNKIAL"
11481. .18476
/gene="L"
11481. .18476
/gene="L"
/note="major component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="L protein"
/protein_id="AAR85466.1"
/db_xref="GI:40388394"
/translation="MQHPTOVPDARLSSPILLDQCLLARSILGLYSHVSHNPKLNCR
IPHHIYRLNSTALTKFLQNCISILVPHSINWDLITSIQDAINHVDDFKYLLPSEL
VKIANDNEFLKAYLNLKILGLDHFVSASARSCBDFSPKNPYMGMLLLVHLSQLAR
RIKQGRSLRNMKFIGTDLFLGFIADFVIFKVPVKTIIRNAVLSQASKPGRLIWRD
QNLTPYLCDDPEIVSAVSECFIMIKDVFIERVNTWEICARAWLEDSQADYPPDLVL
GELVNGQDOI I AMYLEDGFKLILHLEPLVCSCIOQTHGFTPRKYWFQSQMIKSYVDEL
HDNLKLOI SDNKASCAONFIKTI VOAKLTPOYCELP SLOKMGHPVLYNDVALDKV
KHAQSTKLKPKWPFECVFKPIVAKNHSFQSGSYKTTDHLTPYLRLQHVSNV
FPSAEITVOHLMWEIFVSEHPLFSTKIIISDLSIIFKDRATAVNOECWDSVFRSIVLY
NPLVFRQSVQETPQADSLFNQILFEFAEKLIALPSYRNFSLKXELNTRFT
GKPLVRVNOVPTLAALLADGLAKAFPPSNMVMVTEREQKALLHQASWHHNSASIGEN
ALVRGASVETEDNRPDCANA VHYHLGIGLEGLQKLTWCISCAOITVELTKLPKLS
YSPHCVTEDNRPDCANA VHYHLGIGLEGLQKLTWCISCAOITVELTKLPKLS
CYMGDQCITLISLFPIDAPNDYQNEAELNARVABIAITGVSGLFKPBEFTVH
SGFIYFGRKQYLVOLPQSLTMAWCPGLSDSIDDLQGSLSIAGTSFERTGETSTRH
IFPSRISFMSHMLAINLNQHLGFLPGLFNIDISCFKPLTFSEKLLALITPQVLG
LSFLNPKLFYENISDPTLSGLFQNLKALEFLEKEFLPYILI SKRPGLADASDFVWNP
LGLNVPKSEIITFLRQTVRENTITSONRIINSLFHIGSDLEDORVCEWLLSNPVM
SRFADIDFRTSPGKLOVLGVEGTRTILASRTISLATTEGTMKMLRELTRNWKSW
FSYIDALDDDLSELEKFTCTVDVANFLRAYSWSVGLKRLIGATLPLEQEFVKW
INLSLEDQFNLSSDSKSTINLLPYDCKELELEGSDNTELNYVSCALDRKVVQKHS
VNLAWTIGNRAPYIGSRTEDKIGYPPLRVNCPFAALEKEAIEWVSRLLWNTVQGTADRE
KLLIPLLSRVNLVDQTVLNFPLTHYSGNIVHRYNDQYQHSFMANRMSNTSTRAIIS
TNTLGKYGAGGQAAIDSNII FONTINLGAVLIDIALSLAKLSASNVTFRLMLNKCTT
RHVPSEYLYFDKPLVDNLKMDNELVYDNDPLCSGIKGRIGRUSVRSITLSLNVSDI
GSDIPTTAWTILGETTVGSIFSDSESSQSTDPDISSGCTKTFVTHFLVFPVESIFYAFG

ANLIVESLSRIKSLKNSLDLPLTSLISRTNLSHRSLSRILOSTFRHELVLRLAHHI
PLISLMLGSGAGKSSDAVRLELTASYQNFNNFSCIMKMGQSLSLFWLYFPBEGQQ
LKPLKTLQRLSDLLSPDKIKRLIADTCCPIGFSFWYPSKSTRTHYASLANWRD
KANKVNTPFSLINCSPEFSSHTSVSNQVNTSKYIYIPENITAINARTLLINY
GSTALQGMDTKMLSEQLNVCNRPSEIRFKONQKITKHQDQ CERESSESPQMPED
NMOPTAHIHSSSPFOILI KSLDAHEDFASKI ILNSENINNLNTLEYTLNKLTPTR
TEILDTSPLOSSSVLSSTRSLSTRSQASVLYVDCNSIPISIDPGRSMSONOVQ
MLINTYKRDHLACFDSNQFCRFTGVSSMYKLYDLPLPPGKLKXAI CLAESEGGGARL
LLWKETDYILFNTLATDSQEAELISGRVIRPMLYINIDRISLLESRRLLNLNLTIO
IITDNTLMDLSVLIQYLPBDSIDLTDMAETTKDREQLYKTIYINITRSPNPKIS
IIKVFLDYGFTLFLMKNAIQYVGOVLKPKYSSNAKNSWFLTGCGKRRIRLQLODFS
DOVGIFLICAMSRORQAI PVMLKHIENKYPASLHEFELTGLPSPLESSFCHRYTIPF
SEKALPHKQSVYVRQKQHLHSIMLDYENNSPLLDLNRHPICLSIRKITYKNDILK
LNLVIAVEKGRKWSQVLEILPNHVSVCIVHVDHECSCGCEKRLLLKLLDFIRNTKIAEQ
KLLNRVIGYILFFPPFLGSGSURA"

ORIGIN

	Query Match	70.7%;	Score 1301.8;	DB 14;	Length 19113;
	Best Local Similarity	99.8%;	Pred. No. 0;		
	Matches 1303;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC	60		
Db	5941	ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC	6000		
Qy	61	ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT	120		
Db	6001	ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT	6060		
Qy	121	CTCAGAAGACAGAAAGAGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180		
Db	6061	CTCAGAAGACAGAAAGAGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	6120		
Qy	181	GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT	240		
Db	6121	GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT	6180		
Qy	241	GTTGAGTACACAGAGGGGAGGAGGCAAAACATGCTCAATAATAAGTGTAAACGATCCC	300		
Db	6181	GTTGAGTACACAGAGGGGAGGAGGCAAAACATGCTCAATAATAAGTGTAAACGATCCC	6240		
Qy	301	TCTGGAATACTCTTGTTGTAGATCCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	360		
Db	6241	TCTGGAATACTCTTGTTGTAGATCCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	6300		
Qy	361	ACTATCCATCATATTCAAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTATGG	420		
Db	6301	ACTATCCATCATATTCAAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTATGG	6360		
Qy	421	GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGTCTTC	480		
Db	6361	GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGTCTTC	6420		
Qy	481	ACTGAAGGAACATACGACGTCTATGTTGTCATTAAGCAGTGCACAAATGATTTTCTCG	540		
Db	6421	ACTGAAGGAACATACGACGTCTATGTTGTCATTAAGCAGTGCACAAATGATTTTCTCG	6480		
Qy	541	CGGCAAGGACAAAGGTACCGTTCATATGAATCTGACTTCTACTAATAAATATTTGCAAGT	600		
Db	6481	CGGCAAGGACAAAGGTACCGTTCATATGAATCTGACTTCTACTAATAAATATTTGCAAGT	6540		
Qy	601	AGTAAACGGAACGCAAAACGAATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT	660		
Db	6541	AGTAAACGGAACGCAAAACGAATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT	6600		
Qy	661	ACAAGAACCAAAACATGTCCTCCGTCGAAATACCTCCACCACTGCCACACCGCGTCCG	720		
Db	6601	ACAAGAACCAAAACATGTCCTCCGTCGAAATACCTCCACCACTGCCACACCGCGTCCG	6660		
Qy	721	GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCAAGGACCCCAAGC	780		
Db	6661	GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCAAGGACCCCAAGC	6720		

QY 1201 GGGCCGGCCAGACCGCACCAACAGGACCAATGAGCATTTTCCAGTCTCTCCCGCCACC 1260
Db 1198 GGGCCGGCCAGACCGCACCAACAGGACCAATGAGCATTTTCCAGTCTCTCCCGCCACC 1257
QY 1261 CCAGCTCGACTGCACAACTCTTGTATATTTTTCAGAGAAAGCGATCG 1308
Db 1258 CCAGCTCGACTGCACAACTCTTGTATATTTTTCAGAGAAAGCGATCG 1305

RESULT 8
AF005735 2046 bp RNA linear VRL 23-JAN-1998
LOCUS Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein
DEFINITION precursor (GP) gene, complete cds.
ACCESSION AF005735
VERSION
KEYWORDS
SOURCE
ORGANISM Lake Victoria marburgvirus
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PUBMED 9448698
REFERENCE 2 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA

FEATURES
source
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Germany/Marburg/1967/Ratayczak"
/db_xref="taxon:11269"
gene
1..2046
/gene="GP"
CDS
1..2046
/gene="GP"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40460.1"
/db_xref="GI:2459880"
/translation="MKTTCLFLSLILIQIKTLPILEIASNNQPNQVDSVCSGTLQKT
EDVLMGFTLSGQKVDSPLEASKEWAFRTGVPKQNVETGEERAKTCYNI SVTPDSG
KSLLDPPNIRIDYPKCKTIHI QSNPHAGQIALHLAGPELYDRIASTTMYGRVF
TEGNTAAMLVNKHVHMIFESROGQYRHNNLSTNKYTSNNGTQTNDTGCEGALQEY
NSTKQTCAPSKIPGLPTPEIRIAPTSTPTQDNTLNTDNNDEDLITSGSGSEQ
EPTTSDAVTQKGLSTWPPSPQSPFPQEGNNDHSQGTVPKNTNTTAQSPMP
PHNTAISTNTSKNFTLSVSLQNTYDTSATENEQTSAPSKTTLPTGLNLT
AKSTNWKPTTATPNMNGHLSPTSPNPTTQHLVYFRKRS ILWREGDMFPDLG
LNPADIDPPVNTKI IDPSSSSGASAEEDHASPNI SLTILSYFPNINEMTVSGEN
ENDCAELRINSVQEDDLAAGLSWIPFPFGLEGHYTAGLIGNQNVLVRLRLANQT
AKSLELLURVTEFTFSLINRHIDFLLRWGTGCKVLGPDCCICIDELSRNISEQI
DQIKDEQKEGTGWLGGKWTSDWGLVTNLGILLLLLSIAVLIALSCIRIPTYIG"
1..81
/gene="GP"
sig_peptide
82..2043
mat_peptide
/gene="GP"
/product="glycoprotein"
misc_feature
1930..2019
/gene="GP"
/note="encodes transmembrane anchor peptide"

Query Match 62.5%; Score 1151.4; DB 14; Length 2046;

Best Local Similarity 92.6%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTCTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
QY 121 CTCAGAGACAGAGAGCGTCCATCTGTATGGATTTCACTAGTGGGCAAAAAAGTGTCT 180
Db 121 CTCAGAGACAGAGAGATGTCCTCTGTATGGATTTCACTAGTGGGCAAAAAAGTGTCT 180
QY 181 GATTCCCTTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCCTTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
QY 241 GTTGAGTACAGAGGGGGAGAGCCAAACATGCTACAATAAAGTGTACCGATCCC 300
Db 241 GTTGAGTATACAGAGGGGGAGAGCCAAACATGCTACAATAAAGTGTACCGATCCC 300
QY 301 TCTGMAAATCCTTGTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTGMAAATCCTTGTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGATTTTTTCTGTATGATCGATGCTCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGAGATTTTTTCTGTATGATCGATGCTCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
QY 481 ACTGAGGGAACATAGCAGCTATGATGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
Db 481 ACTGAGGGAACATAGCAGCTATGATGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
QY 541 CGGCAAGCAACAGGATCCGTCAATGATCTGACTTCTACTATAATAATTTTGGACAAAGT 600
Db 541 AGGCNAGCAGAGGGGTACCGTCACATGATCTGACTTCTACTATAATAATTTTGGACAAAGT 600
QY 601 AGTAACGAAACGCAAAACGAATGACATGCGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 660
Db 601 AACATGGAACACAAACGAATGACATGCGATGTTTGGCGCTCTTCAAGAATATCAAACTCC 660
QY 661 ACNAGAACCAACATGTCCTCGTCCAAATACCTCCACCACTGCCCCACAGCCCGTCCG 720
Db 661 ACNAGAACCAACATGTCCTCGTCCAAATACCTCCACCACTGCCCCACAGCCCGTCCG 720
QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCGCAACCAACTCAATACCAAGGACCAAGC 780
Db 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCGCAACCAACTCAATACCAAGGACCAAGC 780
QY 781 AGTGTATGAGGACCTCCGCAACATCCGGCTCAGGGTCCGAGAAACGAGAACCCCAACA 840
Db 781 AATGATGATGAGGACCTCATAACTCCGGTTCAGGGTCCGAGAAACGAGAACCCCTATACA 840
QY 841 ACTTCTGATGGGTCCACAGCAAGGGCTTTTCATCAACATGCGCACCCACTCCCTCACC 900
Db 841 ACTTCTGATGGGTCCACAGCAAGGGCTTTTCATCAACATGCGCACCCACTCCCTCACC 900
QY 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACCAACCACTTCCCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACCAAGATGCTGTGACT 960
QY 961 GAACTAGACAAAAATAACAACTGACAAACCGTTCCTGCCCCCTCATAAACAATACCA 1020
Db 961 GAACTAGACAAAAATAACAACTGACAAACCGTTCCTGCCCCCTCATAAACAATACCA 1020
QY 1021 ATCTCTACTTAACACACCTCTCCAAACAACTTTCAGACTCTCTCTGACCATTACAAAC 1080
Db 1021 ATCTCTACTTAACACACCTCTCTCCAAACAACTTTCAGACTCTCTCTGACCATTACAAAC 1080

```
Db 1021 ATCTCTATCAACAACACCTCCAGAACAACTTCAGACACCTCTCTGTATCATCAACAAC 1080
Qy 1081 ACCACCAATGACAAACACACAGACGACAACTCACTGAAATGACAAACAGTCCCCCTCG 1140
Db 1081 ACCACCAATGACAAACACACAGACGACGACGACGACGACGACGACGACGACGACGACG 1140
Qy 1141 ATAAACACCTGCTCCCAACAGGGAATCCACACAGCAAGAGACGACGACGACGACGACG 1200
Db 1141 AAACACACCTGCTCCCAACAGGGAATCTTACACAGCAAGAGACGACGACGACGACG 1200
Qy 1201 GCGCGGCGCACACAGCGCACCAACAGCAAGATGAGCATTTCCAGTCTCCCTCCAC 1260
Db 1201 GCGCGGCGCACACAGCGCACCAATATGACAAATGGCATTTTAAACAGTCTCCCTCCAC 1260
Qy 1261 CCAGCTGCACTGCAACATCTTGATATTTTTCAGAGAAAGCGA 1305
Db 1261 CCACACCGACACACACATCTTGATATTTTTCAGAGAAAGCGA 1305

RESULT 9
LOCUS MAVSPA 2948 bp RNA linear VPL 22-JAN-1996
DEFINITION Marburg virus genomic RNA of GP gene.
ACCESSION X68493
VERSION X68493.1 GI:296960
KEYWORDS gp gene; structural protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2948)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The GP-protein of Marburg virus contains the region similar to the
'immunosuppressive domain' of oncogenic retrovirus P15S proteins
FEBS Lett. 323 (1-2), 183-187 (1993)
JOURNAL MEDLINE 93265932
PUBMED 8495737
REFERENCE 2 (bases 1 to 2948)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and
Netesov,S.V.
TITLE The complete nucleotide sequence of the Popp (1967) strain of
Marburg virus: a comparison with the Musoke (1980) strain
Arch. Virol. 140 (9), 1589-1600 (1995)
JOURNAL MEDLINE 96028047
PUBMED 7487490
REFERENCE 3 (bases 1 to 2948)
AUTHORS Bukreyev,A.
TITLE Direct Submission
Submitted (28-SEP-1992) A.A. Bukreyev, All-Union Inst. of Molecular
Biology, NPO 'Vector', Koltsovo, Novosibirsk region, 633159, USSR
LOCATION/Qualifiers
1. 2948
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="pop"
/db_xref="taxon:11269"
6. 2851
/gene="GP"
(6.17) . . (2841.2851)
/gene="GP"
122. 2167
/gene="GP"
/codon_start=1
/product="structural protein"
/protein_id="CAA48507.1"
/db_xref="GI:296961"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
translation="WKTCLFSLILIQIKTLPILEIASNNPONVDSVCSGTLTQKT
EDVHLMGFTLSQKVADSPLEASKWAFETGVPKNVEYTEGEAKTCYNISVTPSG
KSLILDPTNIRDYPKCKTIHHIQONPHAQALHLGAFELYDRIASTMYGRVF
TGNATAAMLVNKHMIIFRQGGYRHNLTSTNKYMTSNNGTQNDTDCFGALQEY
NSTKNQTCAPSKIPSLPTARPEIKPTSTPTDATTTLNTDPPNDDDELITSGSGSGE
```

ORIGIN

```
Query Match 62.4%; Score 1148.2; DB 14; Length 2948;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 1 ATGAAGACACATGTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 60
Db 122 ATGAAGACACATGTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 181
Qy 61 ATTTTAGAGATAGTAGTAATAACCAACCAAAATGTGGATTGGTATGCTCCGGAAT 120
Db 182 ATTTTAGAGATAGTAGTAATAACCAACCAAAATGTGGATTGGTATGCTCCGGAAT 241
Qy 121 CTCAGAGAGACAGAGAGCGTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 242 CTCAGAGAGACAGAGAGATGTCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 301
Qy 181 GATTCCCCCTTTGGAGGACATCCAAAGCGATGGGCTTTTCAGGACAGAGTGTACCTCCCAAGAT 240
Db 302 GATTCCCCCTTTGGAGGACATCCAAAGCGATGGGCTTTTCAGGACAGAGTGTACCTCCCAAGAT 361
Qy 241 GTTAGATACACAGAGGGGAGAGCAAAACATGCTACATATAAGTGTAAAGATGATGATGCTCC 300
Db 362 GTTAGATACACAGAGGGGAGAGCAAAACATGCTACATATAAGTGTAAAGATGATGATGCTCC 421
Qy 301 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGAGTATCCGAAATGCAAA 360
Db 422 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGAGTATCCGAAATGCAAA 481
Qy 361 ACTATCCATCATATTCAGAGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATG 420
Db 482 ACTATCCATCATATTCAGAGTCAAAACCTCATGCAGAGGGATCGCCCTTCATTATG 541
Qy 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAGTCTTC 480
Db 542 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAGTCTTC 601
Qy 481 ACTGAAGGGAACATAGCAGCTATGATCTCAATAGACAGTGCACAAATGATTTCTCG 540
Db 602 ACTGAAGGGAACATAGCAGCTATGATCTCAATAGACAGTGCACAAATGATTTCTCG 661
Qy 541 CGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGAAGT 600
Db 662 AGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGAAGT 721
Qy 601 AGTAACGGAACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTCAAGATATCAATTT 660
Db 722 AACATGGAACACAAACGAATGACACTGGATGTTTGGCGCTCTTCAAGATATCAATTT 781
Qy 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 720
Db 782 ACAGAGAAATCAAAACATGTGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGTCC 841
Qy 721 GAGATCAAACTCAAGACACCCCACTGATGTCACCAAACTCAATACCAAGGACCAAGC 780
Db 842 GAGATCAAACTCAAGACACCCCACTGATGTCACCACTCAATACCAAGGACCAAGC 901
Qy 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAAACGAGAAACCCACACA 840
Db 902 AATGATGATGAGGACCTCGCAACATCCGGTTCAGGGTCCGAGAAACGAGAAACCCATACA 961
Qy 841 ACTTCTGATGGGTCCACCAAGAGGGTTCATCAACAAATGCCACCACTCCCTCAACA 900
Db 962 ACTTCTGATGGGTCCACCAAGAGGGTTCATCAACAAATGCCACCACTCCCTCAACA 1021
```

QY	901	CAACACGACGCCACACGAGGAGGAACACACAAACCATTTCCACAGATGCTGTGACT	960
Db	1022	CAACCAAGCACGCCACACGAGGAGGAACACACAGACCATTTCCCAAGGTACTGTGACT	1081
QY	961	GAACTAGACAAAATAACACAACTGACACACCGTCCATGCGCCCTCATATAACACTACACACA	1020
Db	1082	GAACCCACAAAACCAACACACAGGACACACCGTCCATGCGCCCTCCACACACACTGCA	1141
QY	1021	ATCTCTACTAACACACCTCTCAAAACACAACTTTCAGCATCTCTCTGACACCATTAACAAAC	1080
Db	1142	ATCTCTACTAACACACCTCTCAAAACACAACTTTCAGCATCTCTCTGATCACTACAAAC	1201
QY	1081	ACCAACCATGACACACACAGGACCAATCACTGGAATGACCAACAGTGCCTCCCTCG	1140
Db	1202	ACCAACCATGACACACAGGACCAATCACTGGAATGACCAACAGTGCCTCCCTCG	1261
QY	1141	ATAACAAACCTGCTCCCAACGGGAATCCACACAGCAAGAGACGACGACGACAAAAA	1200
Db	1262	AAACACACCTGCTCCCAACGGGAATCTTACACAGCAAGAGACGACGACGACGAAA	1321
QY	1201	GCCCCCGCACAAACGGCACCAACACGACCAATGAGCATTTTCCAGTCTCTCCCTCCAC	1260
Db	1322	GCCCCCGCACAAACGGCACCAATGAGCATTTTAAACAGTCTCTCTCCCTCCAC	1381
QY	1261	CCCACTGACGTCGACACATCTTGTATATTTTCAGAGAAACGGA	1305
Db	1382	CCCAACCCGACACACACATCTTGTATATTTTCAGAGAAACGGA	1426
RESULT 10			
MVIRPR			
LOCUS	MVIRPR	19112 bp	RNA linear VRL 22-JAN-1996
DEFINITION	Marburg virus (Popp)	NP, VP35, VP40, GP, VP30, VP24, L genes.	
ACCESSION	Z29337		
VERSION	Z29337.1	GI:450908	
KEYWORDS	GP protein; L protein; NP protein; NP protein; vp24 protein; vp30 protein; vp35 protein; vp40 protein.		
SOURCE	Lake Victoria marburgvirus		
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales;		
REFERENCE	1 (bases 2851 to 5818)		
AUTHORS	Bukreyev, A., Volchkov, V.E., Blinov, V.M. and Netesov, S.V.		
TITLE	The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses		
JOURNAL	FEBs Lett. 322 (1), 41-46 (1993)		
MEDLINE	93245956		
PUBMED	8482365		
REFERENCE	2 (bases 5824 to 8669)		
AUTHORS	Bukreyev, A., Volchkov, V.E., Blinov, V.M. and Netesov, S.V.		
TITLE	The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins		
JOURNAL	FEBs Lett. 323 (1-2), 183-187 (1993)		
MEDLINE	93265932		
PUBMED	8495737		
REFERENCE	3 (bases 1 to 19112)		
AUTHORS	Bukreyev, A., Volchkov, V.E., Blinov, V.M., Dryga, S.A. and Netesov, S.V.		
TITLE	The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain		
JOURNAL	Arch. Virol. 140 (9), 1589-1600 (1995)		
MEDLINE	96028047		
PUBMED	7487490		
REFERENCE	4 (bases 1 to 19112)		
AUTHORS	Bukreyev, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-JAN-1994) Bukreyev A.A., Institute of Molecular Biology, Laboratory of molecular virology, Koltsovo, Novosibirsk region, Russia, 633159		
FEATURES	Location/Qualifiers		
source	1..19112		
	/organism="Lake Victoria marburgvirus"		
	/mol_type="genomic RNA"		

/strain="Popp"	
/db_xref="taxon:11269"	
103..2190	
/gene="NP"	
103..2190	
/gene="NP"	
/function="nucleoprotein"	
/citation=[3]	
/number=1	
/codon_start=1	
/label=NP	
/product="NP protein"	
/protein_id="CAA82536.1"	
/db_xref="GI:450909"	
/db_xref="GOA:P35263"	
/db_xref="UniProt/Swiss-Prot:P35263"	
/translation="MDLSLLELGTKPTAPHRNKKVILFDTHQVSIICNOIIDAINS GIDGLLEGGLTLCVEHYNSDKDFNTSPIAKYLRDAGVDFDVKNAADATPLDVG IPNEPHYSPLIALKLTLESTESQGRIGLFLSCSLPLPLAVGDRASIEKALQVTV GQTFSGLLIVKTVLEFILQKTSGLVTLPKLVRTSKVNEVASFKQALSNLARHGEYA PFARVLNSGINNLEHGLYQLSAIALGVAHTAGSTLAGNVNGOYQQLREAAHDAE KLQRHBOHQIAEDDEERKILEQPHLOKTEITHSOTLAVLSOKREKLAAAEIE NNIVEDQFKQSQNSQSLNDPTPVETVQARPINRPTALPVPDNKRIEHESTEDS SSSSFVDLNDPFDLNEDEDTLDDSVMIPTSTTSREFOGIPAPRQSQDLNNSOGKQ DESTNPKKFLRYQELPPVORDESEYITDQSDIDQSGDNQGVGLPPPLPQAQ KRODPIQHPAVSQDPFGSGIDVNGDILEPIRSPSPSAPOQEDTRAREAYBELSPDFTN YEDQNQWPQVVTTKGRFTFLYPNDLLQTNPPESLITALVEEQNPVSAKELQADWPD MSFDERRHVAMNL"	
2944..3933	
/gene="VP35"	
2944..3933	
/gene="VP35"	
/function="minor component of RNA-dependent RNA-polymerase?"	
/citation=[1]	
/citation=[3]	
/number=2	
/codon_start=1	
/label=VP35	
/product="VP35 protein"	
/protein_id="CAA82537.1"	
/db_xref="GI:450910"	
/db_xref="UniProt/Swiss-Prot:Q03039"	
/translation="MWDSSYMQVSEGLMTGKVPIDQVFGANPSEKIHRRKPKGTG LQCSPLMSKATSDTDIVMDOLIVKTKTLADLLIPINQISDIQSTLNVEITRVHEIER OLHEITPVLMKRTLEAISKGMSEMLAKYDHLVISTGTRTAPAAAFDAYLNEHGVPP QPAIFKDLGVAQAQCSKGTWVNETTDAADKMSKVLSESETEFKPNLSAKDLALLP THLFGNTPPHILAQVLSKIAYKSGKAGFLDAFHQILSEGENAQAALTRLSPFDAP LGVPVPVIRVKNFQTPRCPQKSLRAVPPNPTIDKGWVCVYSSSQGETRALKI"	
4567..5478	
/gene="VP40"	
4567..5478	
/gene="VP40"	
/function="matrix protein?"	
/citation=[1]	
/citation=[3]	
/number=3	
/codon_start=1	
/label=VP40	
/product="VP40 protein"	
/protein_id="CAA82538.1"	
/db_xref="GI:450911"	
/db_xref="GOA:Q03040"	
/db_xref="UniProt/Swiss-Prot:Q03040"	
/translation="MASSSNYNTYMQVLPNPPVADHGANQLIPADLSNOQGITPNVY GDLNDDQKGNVCHAPTLEAIDISATNPTVKGVPKAWLPIGINSFEYFLAHTVAA LLTGSYITQTFHNGQKFPVRVIGTGPAPHLRMLREGNOAFIQNMVPRNFSTNQ TYNLTNLVSQKLPDDAWRPSKDLIGNTMHPAVSIHPLPVLVPTVKKQAVRQHK NPNNGLIAISGILHQLRVEKVPKTSLFRI SLPADMFVSKEGMMKKRGENSEPVVYFQ APENFPLNGENNRQVVLAYANPTLSAV"	
5940..7985	
/gene="GP"	

Qy	841	ACTTCTGATGGCGTCA	CAAGCAGAGGGCTTT	CATCAACAATG	CACCCACTCCCTCA	CCA	900
Db	6780	ACTTTCAGATGGCGT	CAC	TAAGCAAGGGCTTT	CATCAACAATG	CACCCACTCCCTCA	6839
Qy	901	CAACCAAGCAGCGC	CACAGCAGGAGG	AAACAAACAAC	CAATTC	CCCAAGATGCTGTGACT	960
Db	6840	CAACCAAGCAGCGC	CACAGCAGGAGG	AAACAAACAAC	CAATTC	CCCAAGATGCTGTGACT	6899
Qy	961	GAACCTAGACAAAA	ATAACACAAC	TGCGCAAC	CCGTCCAT	TGCCCCCTCATTAACACTAC	1020
Db	6900	GAACCAACAACAAA	CCAAACAAC	AGGCGACAAC	CGTCCAT	TGCCCCCTCAACACCACTGCA	6959
Qy	1021	ATCTCTACTAACA	CACCTCCAAACA	CAACTTCAG	CACTCTCTG	CACCACTTACAAAAC	1080
Db	6960	ATCTCTACTAACA	CACCTCCAAAGA	CAACTTCAG	CACTCTCTG	TATCACTACAAAAC	7019
Qy	1081	ACCACCAATGACA	CACACAGAC	GACACAAT	CACTGAA	AATGAGCAAAACAGTGC	1140
Db	7020	ACCAACCAAT	TACGACAC	CACAGAC	GACGCACTG	GAATATGACAAACAGTGC	7079
Qy	1141	ATAACAACCCCTG	CTCCAAACGGG	AAATCCCA	CCACAGCA	AAAGAGCACCAGCAGCA	1200
Db	7080	AAAACAACCCCTG	CTCCAAACAGG	AAATCTTAC	CACAGCA	AAAGAGCACAACACGAAA	7139
Qy	1201	GGCCCCGGCACA	CAACGGCACC	AAACACAG	CAAAATGAG	CAATTCACAGTCTCTCC	1260
Db	7140	GGCCCCCACC	CAACGGCACC	AAATATGAC	AAATGGGCA	TTTAAACAGTCTCTCC	7199
Qy	1261	CCGAGCTCGACTG	CACACACATCT	TGTATATTT	TCAGAGAAG	CGCA	1305
Db	7200	CCCAACCGGAC	CACACACATCT	TGTATATTT	TCAGAGAAG	CGCA	7244
RESULT 11							
AF005733							
LOCUS	AF005733	2046 bp	RNA	linear	VRL 23-JAN-1998		
DEFINITION	Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin glycoprotein precursor (GP) gene, complete cds.						
ACCESSION	AF005733						
VERSION							
KEYWORDS	AF005733.1	GI:2459875					
SOURCE	Lake Victoria marburgvirus						
ORGANISM	Lake Victoria marburgvirus						
REFERENCE	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.						
AUTHORS	Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.						
TITLE	Variation in the glycoprotein and VP35 genes of Marburg virus strains						
JOURNAL	Virology 240 (1), 138-146 (1998)						
MEDLINE	98110148						
PUBMED	9448698						
REFERENCE	2 (bases 1 to 2046)						
AUTHORS	Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.						
TITLE	Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room S8611, Mail Stop G14, Atlanta, GA 30333, USA						
JOURNAL	Location/Qualifiers						
FEATURES	source						
gene	1. .2046						
CDS	1. .2046						
	/organism="Lake Victoria marburgvirus"						
	/mol_type="genomic RNA"						
	/strain="M/S.Africa/Johannesburg/1975/Ozolin"						
	/db_xref="taxon:11269"						
	1. .2046						
	/gene="Gp"						
	1. .2046						
	/gene="Gp"						
	/note="virion surface glycoprotein"						
	/codon_start=1						
	/product="glycoprotein precursor"						
	/protein_id="AAC40458.1"						

```

/db_xref="GI:2459876"
/translation="MRTTCLFSLILILQIKTLPILIEIASNDQPNVDVSCSTQLQKQ
EDVHLMGFTLSGQKVADSPLEASKRWAFRTGVPKKNVEYTEGEAKTCYNISVYDSG
KSLLLDPNTVRYDPKCTIHHIQGQNQAQIALHGWAFFLYDRIASNTMTYRGKVF
TEGNIAMLVNKTVHKMI FSRQOGYRHNLTSTNKVYSSNGTOTNDTCLGTALHEY
NSTENQTCAPSPTPPPTAPRIKPTSPDTATRLNTNPNSDDEDLFTSGSGSGEO
ELYTSDAVTQKLSSTMPPTSPQCTFOCGGNNTNHSODATDELDTNTNTTAQPTPT
SHNTTITNTNTSKNLSLSPAPNTNPTQSMVTENEKTSAPKATLPTPTENPTT
EKSTNTNTKPTLEPTNTNGHTSPSTNPTTHLYIFRRKRSILWREGDMFPFLDG
LINAPIDFDPVNTKTIIDFESSSGASBEDQHASNISLTSYLSHTSNTYASGEB
ENDCDAELRIWSVOREDDLAAGLSWIPLFGPIEGLTAGLIIKNNNLVCLRLRLANT
AKSLELLRVTTPEERTFSLINWIDELLTRMGCTCKVLGPDCCI GIEDLSRNI SBOI
DQIKKDEQKRGTWGLGKRWTSDMGVTNLGLILLLSIAVLIALSCICRIFTKYIG"
1..81
/sig_peptide /gene="Gp"
mat_peptide /gene="Gp"
misc_feature /product="glycoprotein"
1930..2019 /gene="Gp"
/note="encodes transmembrane anchor peptide"

ORIGIN
Query Match          59.7%; Score 1098.6; DB 14; Length 2046;
Best Local Similarity 50.1%; Pred. No. 0;
Matches 1176; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTTCCCTTATCAGTCTTATCTTAATTCACAGGGACAAAAATCTCCCC 60
DB 1 ATGAGGACTACATGCCCTCTTTATCAGTCTTATCTTAATTCACAGGGATAAAACTCTCCCT 60

QY 61 ATTTAGAGTAGCTAGTAAATAATCAACCCAAAAATGTGGATTCGGTATGCTCCGGAAC 120
DB 61 ATTTTGAGATGCCAGTAACGATCAACCCAAAAATGTGGATTCGGTATGCTCCGGAAC 120

QY 121 CTCAGNAGACAGAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCAGAAAAACAGAAGATGTCCATCTGTATGGGATTTACACTGAGCGGGCAGAAAGTTGCT 180

QY 181 GATTCCCTTTTGAGGATCATCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT 240
DB 181 GATTCCCTTTTGAGGATCAACAGGATGGCTTTTCAGGACAGGTGTACCTCTTAAAGAT 240

QY 241 GTTGAGTACACAGAGGGGAGGAAGCAAAACATGCTACAATATAAGTGTAAACGGATCCC 300
DB 241 GTTGAGTACACGGAAGGGAAGAAGCAAAACATGCTACAATATAAGTGTAAACGGATCCC 300

QY 301 TCTGGAAAAATCCTTGTGTAGATCCTCTTACCAACATCCGCTGACTATCGAAATGCAAA 360
DB 301 TCTGGAAAAATCCTTGTGTAGATCCTCTTACCAACATCCGCTGACTATCGAAATGCAAA 360

QY 361 ACTATCCATCATATTCAAGGTCAAAACCCCTCATGCGAGGGATCGCCCTTCATTTATGG 420
DB 361 ACTATCCATCATATTCAAGGTCAAAACCCCTCATGCGAGGGATCGCCCTTCATTTATGG 420

QY 421 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAACAATGTACCGAGGCAAGTCTTC 480
DB 421 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAACAATGTACCGAGGCAAGTCTTC 480

QY 481 ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGAAGTGCACAAAAATGATTTCTCG 540
DB 481 ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGAAGTGCACAAAAATGATTTCTCG 540

QY 541 CGGCAGGACAAAGGTTACCGTCAATGAATCTGACTTCTACTAATTAATTCGACAGT 600
DB 541 AGGCAGGACAAAGGTTACCGTCAATGAATCTGACTTCTACTAATTAATTCGACAGT 600

QY 601 AGTAAACGGAACGAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATTCT 660
DB 601 AGCAACGGAACGCAAAACAATGACACTGGATGTTTCGGCTACTCTTACGAGTACAATTCT 660

QY 661 ACAAGAACCAAAATGCTGCTCCGTCCAAATACCTCCACACTGCCACAGCCCGTCG 720

```



```
Db 661 ACGAAGAACCAAAACATGTCTCGGTCTTAAACACACCCCCACACCCGCCCGTCCG 720
Qy 721 GAGATCAAACTCACAAGACACCCCAACTGATGCACCAAACTCAATACCAAGACCAAGC 780
Db 721 GAAATCAAAACCAACGACCCCAACCGATGCGATGCTAGCTCAACACCAACCAAC 780
Qy 781 AGTGATGATGAGGACCTCCAAACATCCGCTCAGGTCGGGAGACGAGAACCCACACA 840
Db 781 AGTGATGATGAGGATCTCAACACATCCGCTCGGGTCTGGGGAACAGAACTCTATACG 840
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTCATCAACATGCGCCACTCCCTCACC 900
Db 841 ACTTCTGATGCGGTCACTAAGCAAGGGCTTTCATCAACATGCGCCACTCCCTCACC 900
Qy 901 CAACCAAGCAGCCCAACAGCAGAGGAAACAACAACCACTTCCCAAGATCTGTGACT 960
Db 901 CAACCAAGCAGCCCAACAGCAGAGGAAACAACAACCACTTCCCAAGCAGCTTCACT 960
Qy 961 GAACTAGACAAAATAACAACACTGCGACACCGTCCATGCCCTCCATTAACACTACCA 1020
Db 961 GAACTTGTATAACCAACCAACAACTGCACAAACCGCCCGCTCCCAACACCAACCA 1020
Qy 1021 ATCTCTACTAACAACACTTCCAAACCAACACTTCAGCACTCTCTGACCACTTACAAAC 1080
Db 1021 ATCTCTACCAACAACACTTCCAAACCAACACTTCAGCACTCTCTGACCACTTACAAAC 1080
Qy 1081 ACCACCAATGACACACACAGACGACCAATCACTGAAATGACAAACCACTGCCCTCCG 1140
Db 1081 ACCACCAATGACACACACAGACGACCAATCACTGAAATGACAAACCACTGCCCTCCG 1140
Qy 1141 ATAACAACCTGCTTCCAAACCGGAAATCCCAACAGCAGCAAGAGCAGCAGCAAAAA 1200
Db 1141 AAAGCAACCTGCTTCCAAACAGAAATCTACCAAGAAAGAGCAGCAGCAAAAA 1200
Qy 1201 GGGCCGCGCAACAGGCAACAAACAGCAAAATAGCATTTCACAGTCTCTCCCTCCACC 1260
Db 1201 AGCCCCACCACTTGGAAACAAATACAAACAAATAGCATTTCACAGTCTCTCTCCACC 1260
Qy 1261 CCAGCTCCAGTCGACAACTCTGTATATTTTCAGAAAGACGA 1305
Db 1261 CCCAACTCGACTACACACATCTTATATATTTTCAGAGGAAACGA 1305
```

```
RESULT 12
LOCUS AY358025
DEFINITION Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin, complete
sequence.
ACCESSION AY358025
VERSION AY358025.2 GI:34304751
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES: sRNA negative-strand viruses; Mononegavirales;
Filooviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 19151)
AUTHORS Bowen,M.D., Thurman,K., Minor,E., Ibrahim,M.S., Meyer,R.F.,
Malfatti,S.A., Do,L.H., Smith,K.L., McCready,P.M. and Chain,P.S.G.
Direct Submission
TITLE Submitted (01-AUG-2003) Viral Sequencing Group, Lawrence Livermore
NATIONAL Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
COMMENT On Aug 27, 2003 this sequence version replaced gi:33868618.
FEATURES
source
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/S.Africa/Johannesburg/1975/Ozolin"
/db_xref="taxon:11269"
/notes="cRNA sequence"
primer_bind 1..19
gene 123..2210
/genes="NP"
CDS 123..2210
```

```
/gene="NP"
/notes="nucleoprotein"
/codon_start=1
/product="NP protein"
/protein_id="AAQ55255.1"
/db_xref="GI:33868619"
/translation="MDLHSLLELGTGTAPAHVRNKKVILFDTNHQVSI CNQIIDAINS
GIDLDLLEGLLTCVHEYNSDKFNTSPIAKYLRDAGYEFVKNADATFLDV
IPBEPHSPILIALKLTLESTESQGRIGLFLSFLPLKLVGDRASIEKALQVTV
HOEGIVTYPNHLLTGHMKVIFGLIRSSFILKFLVHOGVNLVTGHDAYSII NSV
GOTRFGLLIVKTLERTLOKDSQVTLHPLVRTSKVNEVASFKQALSNLARHGEYA
PFRVRLNLSGINNLEHGLYFQLSAIALGAVATAHGSTLAGVNVGEQYQLREAAHAEV
KLQRHHEHQIAIAEDDEERKILEQFHLQKTEITHSQTLAVLSQRKELKARAAEIE
NIIVEDQGFQSQNRVSQFLNDPTPEVTVOARVNNRPTALPPVDDKIEHSTEDS
SSSFSVDLNDPALLNEDETDLSVMIPSTTSREFQGI PESQSGDOLDNSSQKQE
DEQNPILKQFLRYQELPPVOEDSEYVTDQESI DQPGSDNEQGVLDLPPPLYTQE
KROPIQHPASSQDPFGSIGDVNGDILLEPIRSPSSPAQEDTFRARAYELSPDFTN
YEDQNMQRQVVTVKGTFLYPNDLQTSPPESLVTALVEEYQNPVSAKELQADWPD
MSFDRHRHVAANL"
/genes="VP35"
/notes="putative minor component of RNA-dependent RNA-
polymerase"
/codon_start=1
/product="VP35 protein"
/protein_id="AAQ55256.1"
/db_xref="GI:33868620"
/translation="MWDSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRPKGTG
LQSPCLMSKATSTDDIWDQLIVKTLADLILPINQISDIQSTLSEVTTKRVHIEIR
OQHEITPVKQVLEAIKSGMSEMLAKYDLHVLVISTGRTTAPAAAFDAYLNHGVPPP
QALFPGDLQAQACSKGTVMYKATDAADKMSKVLSESEFSEKPNLSAKLALLLF
THLPNQTTPHILAQVLSKIAYKSGKCAFIDAPHQILISEGENAAALTRSLTDFAF
LGVVPVIRVKNQTVPRPCSKSLRAVPPNTIDKGVYCVYSSEQGETRALKI"
/genes="VP40"
/notes="putative matrix protein"
/codon_start=1
/product="VP40 protein"
/protein_id="AAQ55257.1"
/db_xref="GI:33868621"
/translation="MASSNNTYMOYLNPPPYADHGANOILIPADQLSNHQHGITPNV
GDNLDDQFGNVCHAFLEAI IDISAVNERTVKGVPAWLPLGIMNSFEYPLAHTVAA
LLTGSYITQTHNGQKQFVRNRLGTGIPAPHLMLREGNOAFIQNMVLPKNSYQK
TYNLTNVLVSQKLPDDAWRPSKOKLIGNTHHPAVSIHPNLPPI VLPVVKQAQYRQK
APENFPLNGFNNRQVVLAYANPTLSAI"
/genes="GP"
/notes="envelope transmembrane glycoprotein"
/codon_start=1
/product="GP protein"
/protein_id="AAQ55258.1"
/db_xref="GI:33868622"
/translation="MRTTCLFSLILIGIKTLPLEIASNDQPNQVDSVCSGTQKLT
EDVHLMFTLSGQKVDASLEASRWAFRTGVPKNVYTEGEAACTCYNISVTPSG
KSLIDPPTNVDRYDKCTIHHIQGNPHAQIALHGAFFLYDRIASTTMYRGKVF
TEGNTAAMVKNKTHKMI FSRQGGYRHMLTSTNKYTSNGTNDTCGCTGLHEY
NSTKNOTCAPSKTPEPPPTARPEIKPTSTPTDATLNTNPSNDEDLTTSGSGEQE
ELYTSDATVTKQLSSLTWPPTSPQPGTQGGNNNHSQDATTELDNTNTTAQPTP
SHNTTISTNNTSKHNLSTLSPONTNPNQSMATENEKTSAPKATLPTEPNTT
KSNNTKSPPTMEPTTNGHTSPSTPNSTQHLIYFRKRSILMRGDMPPFLDG
LINAPIDPVPNTKTIIDESSSSGASTEQHASSNLSLTLSYLRPNTAYSGEN
ENDCDAELR IWSVQEDDLAAGLWIPFPGPIEGLYTAGLIKQNNLVCRRLANQI
AKSLELLRVTTERTFSLINRHAIDFLLTWGCTCKVLGPDCCIIDLSNISBQI
DQIKDEQKGTGWLGGKRWTSWGLVNLNLGILLLSIAVLIALSCICRIFTKYIG"
/genes="VP30"
```


CDS	8888..9733 /gene="VP30" /codon_start=1 /product="VP30 protein" /protein_id="AAQ55259.1" /db_xref="GI:33868623" /translation="MQQPRGRSRVTHQAIPISVHETOLPSKPNVTHHPRASWSST RSSTESSPTNHPRARPESTNLSKPPPPKMDCKNWKIGIPCTDLTCNRDHDLDNLT NRELLLMARWMLPNTKVPSPQDCGSPSISKLSKDKQBOTQDLVLENLGHILNY LHRSIGLKLSTLAALSLTACAGIKRNSLINTMTSLHNENLPODQNGVAKIT TGILHDKGQFEAALWQDKRSISLFVQAALYVNNNIPCESSISVQASVYDFILPOS QKGGQ"	primer_bind ORIGIN Query Match Best Local Similarity 59.7%; Score 1098.6; DB 14; Length 19151; Matches 1176; Conservative 0; Pred. No. 0; Mismatches 129; Indels 0; Gaps 0;	INLVKAVEKGNWSOLVETLPNMHVSVCIVHVDHECFGCEKRLLLKLDPIRNTKIAEQ KLNLRVIGYILFFPFLGSKRSLSRA" complement(19133..19151)
gene	10225..10987 /gene="VP24" 10226..10987 /gene="VP24" /codon_start=1 /product="VP24 protein" /protein_id="AAQ55260.1" /db_xref="GI:33868624" /translation="MAELSTRYNLPANVTEKSIINDLNSTARWIKBPSVGGVTVKMG VFVHPNTGMTLLHLKSNFVVPWQQRNLFSLFKNPKSTIIIEPFLALRILGLVAL KQELQQLSILPGRSIVHMLSEWLLLEVTSAIHISPNLGLIYLTSDMFKILMAGVNF ENKMTLAVNDHGKPSIEIKLTQOIIITRVNMGFLVEVRRIDIEPCCGETILSES VVFGLVAEVLREHSQMEKQPLNLTQTMNSKIAL" 11500..18495 /gene="L" 11500..18495 /gene="L"	Qy 1 ATGAAGACCACATCTTTCTTATCAGTCTTATCTTAATTCAGGAGCAGAAAAATCTCCCC 60 Db 5960 ATGAGGACTACATGCTCTTTATCAGTCTTATCTTAATTCAGGAGCAGAAAAATCTCCCC 6019 Qy 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120 Db 6020 ATTTTGAGATAGGACGATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 6079 Qy 121 CTCAGAAAGACAGAAAGACGCTCATCTGATGGATTTCACACTGAGTGGGCAAAAATGTTCT 180 Db 6080 CTCAGAAAACAGAAAGATGTCATCTGATGGATTTCACACTGAGGCGCAGAAATGTTCT 6139 Qy 181 GATTCCCTTTGGAGGCATCAACGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240 Db 6140 GATTCCCTTTGGAGGCATCAACGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 6199 Qy 241 GTTAGATACACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGTAAAGTAACTCCC 300 Db 6200 GTTAGATACACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGTAAAGTAACTCCC 6259 Qy 301 TCTCGAAAATCTCTTGTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360 Db 6260 TCTCGAAAATCTCTTGTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 6319 Qy 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATATG 420 Db 6320 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATATG 6379 Qy 421 GGAGCATTTTCTGTATGATCGCATTCCTCCACCAATATGTACCGAGGCAAGTCTTTC 480 Db 6380 GGAGCATTTTCTGTATGATCGCATTCCTCCACCAATATGTACCGAGGCAAGTCTTTC 6439 Qy 481 ACTCAAGGACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTTCG 540 Db 6440 ACTCAAGGACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTTCG 6499 Qy 541 CGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAAGT 600 Db 6500 AGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAAGT 6559 Qy 601 AGTAAACGGAACGCAAAAGATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 660 Db 6560 AGCAACGGAACGCAAAAGATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 6619 Qy 661 ACAAGGACCAAAATGCTGCTCCGTCGAAATATCTCCACCATCTCCACACGCGCGCTCG 720 Db 6620 ACGAAGACCAAAATGCTGCTCCGTCGAAATATCTCCACCATCTCCACACGCGCGCTCG 6679 Qy 721 GAGATCAAACTCACAAAGCAGCCCAAACTGATGCCCAAACTCAATACACGAGCCCAAGC 780 Db 6680 GAATCAAACTCACAAAGCAGCCCAAACTGATGCCCAAACTCAATACACGAGCCCAAGC 6739 Qy 781 AGTATGATAGGACCTCGCAACATCTCGGCTCAGGGTCCGAGGAAACGAGAAACCCCAACA 840 Db 6740 AGTATGATAGGACCTCGCAACATCTCGGCTCAGGGTCCGAGGAAACGAGAAACCCCAACA 6799 Qy 841 ACTTCTGATGCGGTCAACAAGCAGGGCTTTCATCAAAATGCGACCTCCCTCCCAACA 900 Db 6800 ACTTCTGATGCGGTCAACAAGCAGGGCTTTCATCAAAATGCGACCTCCCTCCCAACA 6859 Qy 901 CAACCAAGCAGCCACAGAGGAGGAAACAAACAAACCAATTCACAGAGTGTGTGACT 960 Db 6860 CAACCAAGCAGCCACAGAGGAGGAAACAAACAAACCAATTCACAGAGTGTGTGACT 6919	Qy 1 ATGAAGACCACATCTTTCTTATCAGTCTTATCTTAATTCAGGAGCAGAAAAATCTCCCC 60 Db 5960 ATGAGGACTACATGCTCTTTATCAGTCTTATCTTAATTCAGGAGCAGAAAAATCTCCCC 6019 Qy 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120 Db 6020 ATTTTGAGATAGGACGATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 6079 Qy 121 CTCAGAAAGACAGAAAGACGCTCATCTGATGGATTTCACACTGAGTGGGCAAAAATGTTCT 180 Db 6080 CTCAGAAAACAGAAAGATGTCATCTGATGGATTTCACACTGAGGCGCAGAAATGTTCT 6139 Qy 181 GATTCCCTTTGGAGGCATCAACGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240 Db 6140 GATTCCCTTTGGAGGCATCAACGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 6199 Qy 241 GTTAGATACACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGTAAAGTAACTCCC 300 Db 6200 GTTAGATACACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGTAAAGTAACTCCC 6259 Qy 301 TCTCGAAAATCTCTTGTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360 Db 6260 TCTCGAAAATCTCTTGTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 6319 Qy 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATATG 420 Db 6320 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATATG 6379 Qy 421 GGAGCATTTTCTGTATGATCGCATTCCTCCACCAATATGTACCGAGGCAAGTCTTTC 480 Db 6380 GGAGCATTTTCTGTATGATCGCATTCCTCCACCAATATGTACCGAGGCAAGTCTTTC 6439 Qy 481 ACTCAAGGACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTTCG 540 Db 6440 ACTCAAGGACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTTCG 6499 Qy 541 CGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAAGT 600 Db 6500 AGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAAGT 6559 Qy 601 AGTAAACGGAACGCAAAAGATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 660 Db 6560 AGCAACGGAACGCAAAAGATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 6619 Qy 661 ACAAGGACCAAAATGCTGCTCCGTCGAAATATCTCCACCATCTCCACACGCGCGCTCG 720 Db 6620 ACGAAGACCAAAATGCTGCTCCGTCGAAATATCTCCACCATCTCCACACGCGCGCTCG 6679 Qy 721 GAGATCAAACTCACAAAGCAGCCCAAACTGATGCCCAAACTCAATACACGAGCCCAAGC 780 Db 6680 GAATCAAACTCACAAAGCAGCCCAAACTGATGCCCAAACTCAATACACGAGCCCAAGC 6739 Qy 781 AGTATGATAGGACCTCGCAACATCTCGGCTCAGGGTCCGAGGAAACGAGAAACCCCAACA 840 Db 6740 AGTATGATAGGACCTCGCAACATCTCGGCTCAGGGTCCGAGGAAACGAGAAACCCCAACA 6799 Qy 841 ACTTCTGATGCGGTCAACAAGCAGGGCTTTCATCAAAATGCGACCTCCCTCCCAACA 900 Db 6800 ACTTCTGATGCGGTCAACAAGCAGGGCTTTCATCAAAATGCGACCTCCCTCCCAACA 6859 Qy 901 CAACCAAGCAGCCACAGAGGAGGAAACAAACAAACCAATTCACAGAGTGTGTGACT 960 Db 6860 CAACCAAGCAGCCACAGAGGAGGAAACAAACAAACCAATTCACAGAGTGTGTGACT 6919

QY 961 GAACTAGACAAAATAACAACTGCAAAACCGTCCATGCGCCCTCATATAACACTACCA 1020
Db 6920 GAACTTGATAACCAACCAACAACTGCAAAACCGCCCAAGCCCTCCCAACACCAACCA 6979
QY 1021 ATCTCTACTAACCAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAA 1080
Db 6980 ATCTCCACCAACCAACCTCCAAACCAACCTTCAGCACTCTCTCTGCACTTACAA 7039
QY 1081 ACCACCAATGACAAACACACAGAGCAACATCACTGAAAATGAGCAAAACAGTCCCTCG 1140
Db 7040 ACCACCAATGACAAACACAGAGCAACATCACTGAAAATGAGCAAAACAGTCCCTCG 7099
QY 1141 ATACAAACCTCCCTCCAAACCGGAATCCCAACACAGCAAGAGAGAGAGAGAGAGAG 1200
Db 7100 AAAGCAACCTCCCTCCAAACCGGAATCCCAACACAGCAAGAGAGAGAGAGAGAG 7159
QY 1201 GGGCCGCGCACAAACCGGACCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 7160 AGCCCGCACCAACAGGAGCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7219
QY 1261 CCAGCTCGACTGCAACACATCTTGATATATTTTCAAGAAAGAGCGA 1305
Db 7220 CCNACTCGACTGCAACACATCTTATATATTTTCAAGAAAGAGCGA 7264

RESULT 13
AX717719
LOCUS AX717719 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 7 from Patent WO02079239.
ACCESSION AX717719
VERSION AX717719.1 GI:29890728
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Grogan,C.C.; Hevey,M.C. and Schmaljohn,A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 7 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
i .2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Raven Glycoprotein 1 and Marburg virus strain Musoke
Glycoprotein 2"

ORIGIN
Query Match 42.4%; Score 780; DB 6; Length 2046;
Best Local Similarity 74.8%; Pred. No. 6.7e-225;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATTAATTCAGAGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCAATATTTCTGATTAGTCTCAITTTAATCCAAAGTATAAAACTCTCCT 60
QY 61 ATTTTAGATAGTAGTATTAATTAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAATTTGCTAGTAAACGCCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAAC 120
QY 121 CTCGAGACAGAGACGCTCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGACAGAGAGATGTTTATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTTGAGGCATCCAAAGCGATGGCTTTTCAGGACAGGTGTACTCCCAAGAT 240
Db 181 GATTCCCTTTTGAGGCATCTAAACGATGGCTTTTCAGGACAGGTGTCTCTCCCAAGAAC 240
QY 241 GTTGAGTACAGAGGGGAGAGAGCAACATGCTACAAATTAAGTGTACAGGATCCC 300
Db 241 GTTGAGTACAGAGGAGAGAGCAACATGTTTACAAATATAAGTGTACAGACCCCT 300

QY 301 TCTGAAAAATCCTTGTCTGTAGATCCTCTTACCACAAATTCCTGTACTATCCGAAATGCAAA 360
Db 301 TCTGAAAAATCCTTGTCTGTAGATCCTCTTACCACAAATTCCTGTACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTCAGGTCAAAACCTTCATGCAAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTCATCATATTCAGGTCAAAACCTTCATGCAAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGGCATTTTTTCTGTATGATCGGTTCCTCTACAAATGTACCGAGGCAAAAGTCTTC 480
QY 481 ACTGAAGGGAACATAGCAGCTATGTTGTCAATAGCAGTGCACAAATGATTTTCTCG 540
Db 481 ACTGAAGGGAATATAGCAGCTATGTTGTCAATAGCAGTGCACAAATGATTTTCTCT 540
QY 541 CGGCAAGGCAAGGGTACCGTCATATGATCTGACTTCTACTAATAATATTTGCAAGT 600
Db 541 AGGCAAGGCAAGGGTATTCGTCAATGAACCTTGACCTCCCAATAATATTTGGAACAGC 600
QY 601 AGTAAAGGGAACGAAACGAAATGACACTGATGTTTTCGGCGCTCTTCAAGAAATACAA 660
Db 601 AGCAATGAACGCAAGAAATGATACGGGATGTTTGGCATCTCTCAAGAAATACAA 660
QY 661 ACAAGGAACCAACATGTTCTCCGTCCAAATACTCCACCACTGCCCAAGCCCGTCCG 720
Db 661 ACAAGCAATCAACATGCTCCCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 720
QY 721 GAGATCAAACTCACAGCAACCCCACTGATGCCCAAACTCAATACCAAGGACCAAGC 780
Db 721 AGCATTCACCTACAAAATACCTCAAAATTAATATCTGCTAAATCTGGAACTATGAACCC 780
QY 781 AGTGATGATGAGGACCTCGCAACATCGGCTCAGGTCGGAGAACAGAACCCCAACA 840
Db 781 AGCGACGATGAGGACCTTATGTTTCGGCTCAGGATCTGGAGAACAGGCGGCCCAACA 840
QY 841 ACTTCTGTATGCGGTCCCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCA 900
Db 841 ACTCTTAATGATGACTCAACAGAAACAAATCGTCAACAAATATGTTCCACTCTCTCACTA 900
QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
Db 901 CATCCAGCACCTCAACACATGAGCAAAACAGTACGAACTCTCTCCCGACATGCTGTAACT 960
QY 961 GAACTAGACAAAATAACAACTGCAAAACCGTCCATGCGCCCTCATATAACACTACCA 1020
Db 961 GAGCAATGGAACCGACCCCAACCAACACAGCAAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 ATCTCTACTAACCAACCTCCAAACCAAACTTCAGCACTCTCTCTGCACTTACAA 1080
Db 1021 ACTCCACCTTATACACTCTCAAGTACAACTCAGTACTCTCTCTCCCTCCAACCGCAAC 1080
QY 1081 ACCACCAATGACAAACAGAGAGCAACATCACTGAAAATGAGCAAAACAGTGGCCCTCG 1140
Db 1081 ATCCACCAATATGATATACAAACGTAAGTACGAGAAAGCAAAACCAATGCTCAGTTG 1140
QY 1141 ATAAACACCTTCCTCCAAACGGGAAATCCCAACAGCAAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 AACCAACTCTAGATCCAAAGAGAAATCCCAACAGGACAGACCAACAGCAACAC 1200
QY 1201 GGGCCCGCACAAACCGGACCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 AACATCATCATGAGCATCAGATATATAAAGCAAAACCCCAATAATTTCTCTCGGAT 1260
QY 1261 CCAGCTCGACTGCAACACATCTTGATATATTTTCAAGAAAGAGCGTGC 1308
Db 1261 TCTAGTCCGACAAACCGCCCTCTATATATTTTGAAGAAAGAGCGATCG 1308

RESULT 14
AX717725
LOCUS

AX717725 2046 bp DNA linear PAT 15-APR-2003

DEFINITION Sequence 13 from Patent WO02079239.
ACCESSION AX717725
VERSION AX717725.1 GI:29890731
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
REFERENCE Viruses; ssRNA negative-strand viruses; Mononegavirales;
AUTHORS Filoviridae; Marburg-like viruses.
TITLES 1
Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.
Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 13 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
source
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="unassigned DNA"
/strain="Raven"
/db_xref="taxon:11269"
/note="chimeric molecule between Marburg virus
Glycoprotein 1 and Marburg virus Glycoprotein 2"

ORIGIN
Query Match 42.4%; Score 780; DB 6; Length 2046;
Best Local Similarity 74.8%; Pred. No. 6.7e-225;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCTTATCATGCTTATCTTAAATTCAGGGGACAAAAATCTCCCC 60
DB 1 ATGAAGACCATATATTTCTGTATTAGTCTCATTTTAAATCCAAAGTATAAAAATCTCCCT 60
QY 61 ATTTAGATAGCTAGTAAATTAATCAACCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
DB 61 GTTTAGAAATTCGTAGTAACGCCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC 120
QY 121 CTCGAGACAGAGAGCTCCATCTGTGGGATTCACACTGAGTGGGCAAAAGTGTCT 180
DB 121 CTCGAAAGACAGAGATGTTTCATCTGTGGGATTTACACTGAGTGGGCAAAAGTGTCT 180
QY 181 GATTCCTCTTGGAGGATCCAGGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAT 240
DB 181 GATTCCTCTTGGAGGATCTAAACGATGGGCTTTTCAGGACAGGTGTCTCCCAAGAAC 240
QY 241 GTTGAGTACAGAGGGGAGGAGCAAAAATGCTACAATATAGTGTAAAGGATCC 300
DB 241 GTTGAGTATACGGAAGGAGAGCAAAAATGCTACAATATAGTGTAAAGGATCC 300
QY 301 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGTCGATATCCGGAATGCAAA 360
DB 301 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGTCGATATCCGGAATGCAAA 360
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG 420
DB 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGACAGGGATTCGCTTCATTTGTTGG 420
QY 421 GGAGCATTTTTCTGTATGATCGATGCTCCACAAATGATGACGAGGCAAGTCTTC 480
DB 421 CGGGCATTTTTCTGTATGATCGGTGCTCTACAAATGATGACGAGGCAAGTCTTC 480
QY 481 ACTGAAGGAAATATAGCAGCTATGATGTCATATAGACAGTCCCAAAATGATTTTCG 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATGTCATATAGACAGTTCACAGAATGATTTTCT 540
QY 541 CGGCAAGGACAGGGTACCGTCTATGATCTGACTTCTACTAATAATATTTGGCAAGT 600
DB 541 AGGCAAGGACAGGGTATGTCATGACTGACTTCTACTAATAATATTTGGCAAGC 600
QY 601 AGTAACGGAACGCAAAATGATGATGCTGCTCCACAAATGATGACGAGGCAAGTCTTC 660
DB 601 AGCAATGAACGCAAGAAATGATGAGGATGTTTGGCATCTCTCAAGAAATACAACTCC 660
QY 661 ACAAGAACCAACATGCTGCTCCAAATATCTTCAACACTGCTCCACAGCCCGCTCG 720

DB 661 ACAACAAATCAACATGATCCCTCCATCTCTTAAACCTCATCTCCCTGCCACAGTAACCTCG 720
QY 721 GAGATCAAACTCAACAGCAGCCCAACTGATGCCCAAACTCAATACACGAGCAACCAAGC 780
DB 721 AGCATTCACCTACAAATCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGGACGAGGACGAGGACGAGGACG 840
DB 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
QY 841 ACTTCTGATGGGTCAACCAAGGCTTTTCATCAACAAATGCCACCACTCCCTCCACCA 900
DB 841 ACTCTTAATGTAGTACCTGACAGAAACAATCGTCAACATATTTGTCACCTCTTCACTA 900
QY 901 CAACCAAGCAGCCACAGAGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT 960
DB 901 CATCCAGCAGCTCACAACTGAGCAAAACAGTACGAATCTTCCGACATGCTGTAAC 960
QY 961 GAACTAGACAAATTAACAACTGCAACCGTCCATGCTCCCTCCCTCATAACTACCA 1020
DB 961 GAGCACAATGGAACCGACCCCAACCAACCAACGACGACGCTCTCAACAACTAATA 1020
QY 1021 ATCTCTACTAAACACACCTCCAAACACAACTTTCAGCACTCTCTGTGACCACTTACAAAC 1080
DB 1021 ACTCCACCTATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCACCGCAAC 1080
QY 1081 ACCACCAATGACACACAGAGCACAATCACTGAAATAGCAAAACAGTGCCTCCCTCG 1140
DB 1081 ATCCAAATATGATACACAACGTGACAGTACAGAAAGCAAAACCAATGCTCAGTTG 1140
QY 1141 ATACCAACCTGCTCCCAACGGGAAATCCCAACCAAGAGGACGAGGACGAGGACGAGG 1200
DB 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACCAAGAGGACGAGGACGAGGACGAG 1200
QY 1201 GCGCCGCGCACAGCGGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1260
DB 1201 AACATCATCATGACATCATGATATACAGCAACCAACCAACCAACCAACCAACCAAC 1260
QY 1261 CCCAGCTCGACTGACACATCTTTGTATATTTATTTATTTATTTATTTATTTATTTATTT 1308
DB 1261 TCTAGTCGACACACCGGCTCTATATATTTATTTATTTATTTATTTATTTATTTATTT 1308

RESULT 15
AF005734
LOCUS Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein
DEFINITION precursor (GP) gene, complete cds.
ACCESSION AF005734
VERSION AF005734.1 GI:2459877
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
REFERENCE Filoviridae; Marburg-like viruses.
AUTHORS 1 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PUBMED 9448698
REFERENCE 2 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
AUTHORS Direct Submission
TITLE Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
JOURNAL Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES Location/Qualifiers
source
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"

```
/strain="M/Kenya/Kitum Cave/1987/Ravn"
/db_xref="taxon:11269"
1. .2046
/gene="Gp"
1. .2046
/gene="GP"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40459.1"
/db_xref="GI:2459878"
/translation="MKTIPLFLSLILIOIKLPLVLEIASNSQPDVDSVCSGLQKT
EDVHMGTSLSQKQVADSPLEASKRWAFRTGPPKNVEYTEGEAKTCINISVDPFG
KSLLDPPSNIRDYPKCTVHHIOGONPHOAGIALHLWGAFFLYDRVASTMYRGKFV
TEGNTAAIVTKVTRMIFPSRQGYRHMNLTSTNKWTSSTNQDNDTGFGLQEV
NSTNQTPPSLPSLPTVPSIHSTNTQAKSGTWPNSDDDELIMISGSGSEQ
GPHTTLVNVTEQKQSSITLSTPLHPSSTOHEONSTNPSRHAUTEHNGTDPTOPATL
LNNWNTPTYNLTKNLSPTSPPTRNINNDTORLASESEQTNAGLNTLLDPTENPTT
GQDNTSNTIIMTSDITSKPTNSPDSPTTPPIYFRKRSIFWKEGDIFFPDG
LINTIDFDPIPTETIIDESEFNTSNEQHTPPNISLSLPSFPDKNGDYATSGEN
ENDCDAELRINSVQEDDLAAGLSWTFPFPGIEGLYTAGLKNQNNLVCLRLRLANQT
AKSLLELLRVVTEERTFSLINRHAIDFLLTRWGTCKVLGPDCICIGIEDLSKNISEQI
DKIRDEQKEETGWLGGKWTSDWGLTNLGLILLLSIAVLIALSCIRIFTKYIG"
1. .81
sig_peptide
/gene="Gp"
82. .2043
mat_peptide
/gene="GP"
/product="glycoprotein"
1930. .2019
misc_feature
/gene="GP"
/note="encodes transmembrane anchor peptide"

ORIGIN
Query Match
Best Local Similarity 74.7%; Score 777; DB 14; Length 2046;
Matches 975; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATGTTTCTTATAGTCTTATCTTAAATTCAGGGACAAAATCTCCCC 60
DB 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAATCTCCCT 60
QY 61 ATTTTAGATAGTAGTAGTAATAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT 120
DB 61 GTTTTAGAAAATTTGCTAGTAGAACCCAACTCAAGATGTAGATTCAAGTGTGCTCCGGAAC 120
QY 121 CTCCAGAACAGAGACGCTCCATCTGATGGATTACACACTGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCCAAAAGACAGAGATGTTTCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCTCTTTGGAGCATCCAAGCGATGGGCTTTTCAAGACAGGTGTACTCTCCCAAGAT 240
DB 181 GATTCCTCTTTGGAGCATCTAAACGATGGGCTTTTCAAGACAGGTGTCTCTCCCAAGAAC 240
QY 241 GTTGTAGTACAGAGGGGAGAGAGCCAAAACATGCTACAAATATAGTGTACAGGATCCC 300
DB 241 GTTGTAGTATACGGAAGGAGAGAACCAAAACATGTTTACAAATATAAGTGTAAACAGACCC 300
QY 301 TCTGGAATCTCTGCTGTTAGATCCTCTCAACATCCGAGTATCTCCGAAATGCAAA 360
DB 301 TCTGGAATCTCTGCTGTTAGATCCTCTCAACATCCGAGTATCTCCGAAATGTTAA 360
QY 361 ACTATCCCATCATATTCAGGTCAAAACCTCATGCAAGGGATCGCCCTTCATTTATGG 420
DB 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGCAAGGGATTCGCCCTCCATTTG 420
QY 421 GGAGCATTTTCTGTATGATCGCATTCGCTCCACAAAGATGTACCGAGGGAAGTCTTC 480
DB 421 GGGCATTTTCTGTATGATCGCTGCTCTACAAACATGTACCGAGGGAAGTCTTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATTGTATATAGACAGTTCACAGATGATTTTCT 540
```

Search completed: August 1, 2005, 02:43:41
Job time : 7917.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 964.473 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctct.....tttgttttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*:

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1841	100.0	1841	10	ABT13451 DNA encod
2	1308	71.0	2046	10	ABT13452 DNA encod
3	1303.4	70.8	2103	8	ACC57764 Marburg v
4	1303.4	70.8	2164	4	AAD04042 Marburg v
5	1303.4	70.8	7005	10	ACC71551 VRC6702 (
6	1303.4	70.8	7778	10	ACC71550 VRC6701 (
7	1303.4	70.8	8256	10	ACC71552 VRC6710 (
8	1303.4	70.8	11460	3	AAX87211 VRS repli
9	1287.4	69.9	2100	3	AAX51039 Marburg v
10	1274	69.2	2051	10	ABT13454 DNA encod
11	1148.2	62.4	19112	13	ADM48345 Marburg v
12	780	42.4	2046	10	ABT13453 DNA encod
13	780	42.4	2046	10	ABT13456 DNA encod
14	773.8	42.0	2247	4	AAD04043 Marburg v
15	688	37.4	6902	10	ACC71562 VRC6703 (p
16	540	29.3	2039	10	ABT13455 DNA encod
17	527.6	28.7	2224	3	AAX51038 Ebola vir
18	527.6	28.7	7272	2	AAX59393 Plasmid p
19	527.6	28.7	7285	2	AAX89797 DNA of pV
20	527.6	28.7	9106	12	ADF85419 Ebola vir

21	527.6	28.7	18959	13	ADM48337	Adm48337 Zaire ebo
22	527.6	28.7	18959	13	ADM48364	Adm48364 Zaire ebo
23	527.6	28.7	18959	13	ADM48355	Adm48355 Zaire ebo
24	526	28.6	2298	3	AAX87189	Aax87189 Ebola vir
25	526	28.6	2298	4	AAF76953	Aaf76953 Ebola vir
26	526	28.6	2298	12	ADL27460	Adl27460 Nucleotid
27	526	28.6	6467	10	ACC71530	Acc71530 VRC6052 (
28	526	28.6	6624	10	ACC71523	Acc71523 VRC6002 (
29	526	28.6	7154	10	ACC71521	Acc71521 VRC6000 (
30	526	28.6	7188	10	ACC71522	Acc71522 VRC6001 (
31	526	28.6	7272	2	AAX89798	Aax89798 DNA of pV
32	526	28.6	7285	2	AAX59392	Aax59392 Plasmid p
33	526	28.6	8439	10	ACC71548	Acc71548 VRC6603 (
34	525.6	28.5	2172	4	AAD04041	Aad04041 Ebola vir
35	525.6	28.5	6561	10	ACC71524	Acc71524 VRC6003 (
36	451.8	24.5	6914	10	ACC71529	Acc71529 VRC6008 (
37	451.8	24.5	8199	10	ACC71549	Acc71549 VRC6604 (
38	451.8	24.5	10783	10	ACC71545	Acc71545 VRC6600 (
39	420	22.8	7106	10	ACC71528	Acc71528 VRC6007 (
40	336.2	18.3	7044	10	ACC71527	Acc71527 VRC6006 (
41	324.6	17.6	7002	10	ACC71536	Acc71536 VRC6300 (
42	324.6	17.6	7023	10	ACC71541	Acc71541 VRC6351 (
43	324.6	17.6	7036	10	ACC71537	Acc71537 VRC6301 (
44	324.2	17.6	7001	2	AAX89795	Aax89795 DNA of pV
45	324.2	17.6	7003	2	AAX59390	Aax59390 Plasmid p

ALIGNMENTS

RESULT 1

ABT13451
ID ABT13451 standard; DNA; 1841 BP.

XX AC ABT13451;

XX DT 30-JAN-2003 (first entry)

XX DE DNA encoding a chimeric filovirus protein MBGV-GP1/EBOV-GP2.

XX KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX KW immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX OS Unidentified.

XX OS Chimeric.

XX PN WO200279239-A2.

XX PD 10-OCT-2002.

XX PF 31-JAN-2002; 2002WO-US003339.

XX PR 31-JAN-2001; 2001US-0267522P.

XX PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX PI Grogan CC, Hevey MC, Schmaljohn AL;

XX DR WPI; 2003-040651/03.

XX P-PSDB; ABJ18473.

XX PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX PT useful for inducing an immune response against infection of different
XX PT filoviruses, specifically against both Ebola and Marburg viruses.

XX PS Claim 14; Page 70-72; 9app; English.

XX CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX CC of GP2. The chimeric filovirus GP protein is useful for inducing an
XX CC immune response against infection of different filoviruses, specifically
XX CC against both Ebola and Marburg viruses by being used as a vaccine. This
XX CC polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention	
XX	Sequence 1841 BP; 589 A; 500 C; 361 G; 391 T; 0 U; 0 Other;	
Query Match 100.0%; Score 1841; DB 10; Length 1841;		
Best Local Similarity 100.0%; Pred. No. 0;		
Matches 1841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 ATGAAGACACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCCC	60
DB	1 ATGAAGACACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCCC	60
QY	61 ATTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
DB	61 ATTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
QY	121 CTCAGAGACAGAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	180
DB	121 CTCAGAGACAGAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	180
QY	181 GATTCCCTTTTGGAGGATCCAGGGATGGGCTTTTCCAGGACAGGTGTACCTCCCAAGAT	240
DB	181 GATTCCCTTTTGGAGGATCCAGGGATGGGCTTTTCCAGGACAGGTGTACCTCCCAAGAT	240
QY	241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAAATATAGTGTAAAGGATCCC	300
DB	241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAAATATAGTGTAAAGGATCCC	300
QY	301 TCTGGAAATCTTGTGTAGATCTCTTACCAATCCGTGACTATCCGAAATGCAAA	360
DB	301 TCTGGAAATCTTGTGTAGATCTCTTACCAATCCGTGACTATCCGAAATGCAAA	360
QY	361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG	420
DB	361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG	420
QY	421 GGAGCAATTTTCTGATGATCGCATGCTCCACAAATGTACCGAGGCAAAAGTCTTC	480
DB	421 GGAGCAATTTTCTGATGATCGCATGCTCCACAAATGTACCGAGGCAAAAGTCTTC	480
QY	481 ACTGAAGGAACATAGCAGCTATGTTGTCATTAAGACAGTGCACAAATGATTTTCTCG	540
DB	481 ACTGAAGGAACATAGCAGCTATGTTGTCATTAAGACAGTGCACAAATGATTTTCTCG	540
QY	541 CGGCAAGGACAAAGGTPACCGTCATATGAATCTGACTTCTACTAAATATTTGGCAAGT	600
DB	541 CGGCAAGGACAAAGGTPACCGTCATATGAATCTGACTTCTACTAAATATTTGGCAAGT	600
QY	601 AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT	660
DB	601 AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT	660
QY	661 ACAGAGAACCAACATGTTGCTCGTCCAAATACCTCCACCACTGCCACAGCCGCTCG	720
DB	661 ACAGAGAACCAACATGTTGCTCGTCCAAATACCTCCACCACTGCCACAGCCGCTCG	720
QY	721 GAGATCAAACTCACAAGACCCCACTGATGCCACCAAACTCAATACCAAGCAACCAAGC	780
DB	721 GAGATCAAACTCACAAGACCCCACTGATGCCACCAAACTCAATACCAAGCAACCAAGC	780
QY	781 AGTGATGATGAGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACCGGACCCACACA	840
DB	781 AGTGATGATGAGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACCGGACCCACACA	840
QY	841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAAGATGCCACCACTCCCTCACC	900
DB	841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAAGATGCCACCACTCCCTCACC	900
QY	901 CAACCAAGCAGCGCACAGCAAGGAGAAACACACAAACCATTTCCCAAGATGCTGTACT	960
DB	901 CAACCAAGCAGCGCACAGCAAGGAGAAACACACAAACCATTTCCCAAGATGCTGTACT	960
QY	961 GAACTAGACAAAAATAACTGCAACACCGTCCATGCCCTCATTAACACTACCACA	1020
DB	961 GAACTAGACAAAAATAACTGCAACACCGTCCATGCCCTCATTAACACTACCACA	1020
QY	1021 ATCTCTACTTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAAAC	1080
DB	1021 ATCTCTACTTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAAAC	1080
QY	1081 ACCCAATATGACAAACACAGAGACCAATCACTGAAATAGCAAAACAGTGCCCTCG	1140
DB	1081 ACCCAATATGACAAACACAGAGACCAATCACTGAAATAGCAAAACAGTGCCCTCG	1140
QY	1141 ATACAACTCTGCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGCAAGCAAAA	1200
DB	1141 ATACAACTCTGCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGCAAGCAAAA	1200
QY	1201 GGCCCGGCACAAACGGGACCAACACAGCAATGAGCAATTTCCAGTCTCTCCCTCCACC	1260
DB	1201 GGCCCGGCACAAACGGGACCAACACAGCAATGAGCAATTTCCAGTCTCTCCCTCCACC	1260
QY	1261 CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAAGAAAGCGATCGCAATTTGCTCAAT	1320
DB	1261 CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAAGAAAGCGATCGCAATTTGCTCAAT	1320
QY	1321 GCTCAACCCCAATGCAACCCCTTAATTTACTGGAATCTCAGGATGAGAGTCTGCA	1380
DB	1321 GCTCAACCCCAATGCAACCCCTTAATTTACTGGAATCTCAGGATGAGAGTCTGCA	1380
QY	1381 ATCGAGCTGGCTGGATACCATATTTTCGGGCGAGCAGCCGAGGGAAATTTACATAGAGGG	1440
DB	1381 ATCGAGCTGGCTGGATACCATATTTTCGGGCGAGCAGCCGAGGGAAATTTACATAGAGGG	1440
QY	1441 CTAATGCACAATCAAGATGGTTTAACTCTGTGGTTGAGACAGCTGGCCAAACGAGACGACT	1500
DB	1441 CTAATGCACAATCAAGATGGTTTAACTCTGTGGTTGAGACAGCTGGCCAAACGAGACGACT	1500
QY	1501 CAAGCTCTTCAACTGTTCTCGAGAGCCAACTGAGCTACGCACTTTTCAATCTCTCAAC	1560
DB	1501 CAAGCTCTTCAACTGTTCTCGAGAGCCAACTGAGCTACGCACTTTTCAATCTCTCAAC	1560
QY	1561 CGTAAGGCAATTTGATTTCTGCTGAGGATGGGCGGCACATGCCACATTTCTGGGACCG	1620
DB	1561 CGTAAGGCAATTTGATTTCTGCTGAGGATGGGCGGCACATGCCACATTTCTGGGACCG	1620
QY	1621 GACTGCTGTATCGAACCACTGATTTGGACCAAGAACTAAACAGACAAATTTGATCAGATT	1680
DB	1621 GACTGCTGTATCGAACCACTGATTTGGACCAAGAACTAAACAGACAAATTTGATCAGATT	1680
QY	1681 ATTCATGATTTTGTGATATAAAACCTTCCGACCAAGGGGACAAATGCAATTTGGTGACA	1740
DB	1681 ATTCATGATTTTGTGATATAAAACCTTCCGACCAAGGGGACAAATGCAATTTGGTGACA	1740
QY	1741 GGATGGAGCAATGGATACCGGCAAGGTATTGGAGTTACAGGCTTATAATTTGCAGTTATC	1800
DB	1741 GGATGGAGCAATGGATACCGGCAAGGTATTGGAGTTACAGGCTTATAATTTGCAGTTATC	1800
QY	1801 GCTTTATTTCTGTATATGCAAAATTTGTCTTTTGTAGTTGAATTC	1841
DB	1801 GCTTTATTTCTGTATATGCAAAATTTGTCTTTTGTAGTTGAATTC	1841
RESULT 2		
ABT13452		
ID	ABT13452 standard; DNA; 2046 BP.	
XX	AC ABT13452;	
XX	30-JAN-2003 (first entry)	
XX	DNA encoding a chimeric filovirus protein MUS-GP1/RVN-GP2.	
XX	Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;	
XX	immune response; Ebola; Marburg virus; vaccine; gene; ds.	

OS Unidentified.
 OS Chimeric.
 PN WO200279239-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 31-JAN-2002; 2002WO-US003339.
 XX
 PR 31-JAN-2001; 2001US-0267522P.
 XX
 PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
 XX
 PI Grogan CC, Hevey MC, Schmaljohn AL;
 XX
 DR WPI; 2003-040651/03.
 DR P-PSDB; ABJ18474.
 XX
 PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
 PT useful for inducing an immune response against infection of different
 PT filoviruses, specifically against both Ebola and Marburg viruses.
 XX
 PS Claim 16; Page 74-76; 94pp; English.
 XX
 CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
 CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
 CC of GP2. The chimeric filovirus GP protein is useful for inducing an
 CC immune response against infection of different filoviruses, specifically
 CC against both Ebola and Marburg viruses by being used as a vaccine. This
 CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
 CC protein of the invention
 XX
 SQ Sequence 2046 BP; 659 A; 517 C; 402 G; 468 T; 0 U; 0 Other;

Query Match 71.0%; Score 1308; DB 10; Length 2046;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCCC 60
 DB 1 ATGAAGACACATGTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCCC 60

QY 61 ATTTAGAGATGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
 DB 61 ATTTAGAGATGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120

QY 121 CTCAGAGACAGAGAGCTCCATCTGATGGATTCACATCAGTGGGCAAAAAGTTGCT 180
 DB 121 CTCAGAGACAGAGAGCTCCATCTGATGGATTCACATCAGTGGGCAAAAAGTTGCT 180

QY 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
 DB 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240

QY 241 GTTAGATCACAGAGGGGAGAGAGCCAAACATGCTCAATATAGTGTACCGATCCC 300
 DB 241 GTTAGATCACAGAGGGGAGAGAGCCAAACATGCTCAATATAGTGTACCGATCCC 300

QY 301 TCTGGAAATCCTGCTGTAGATCCTCTCAACATCCGCTATCCGAAATGCAAA 360
 DB 301 TCTGGAAATCCTGCTGTAGATCCTCTCAACATCCGCTATCCGAAATGCAAA 360

QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
 DB 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420

QY 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
 DB 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480

QY 481 ACTGAAGGACATAGCAGCTATGATTTCAATAGACAGTGCACAAAATGATTTTCTCG 540
 DB 481 ACTGAAGGACATAGCAGCTATGATTTCAATAGACAGTGCACAAAATGATTTTCTCG 540

QY 541 CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGATCTTCTACTATAATAATATTGACAAGT 600
 DB 541 CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGATCTTCTACTATAATAATATTGACAAGT 600

QY 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGATACAATTTCT 660
 DB 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGATACAATTTCT 660

QY 661 ACAAGAAACAAACATGTGTCTCCGTCCAAATATCTCCACACTGCCCACACAGCCGCTCG 720
 DB 661 ACAAGAAACAAACATGTGTCTCCGTCCAAATATCTCCACACTGCCCACACAGCCGCTCG 720

QY 721 GAGATCAAACTCACAGACCCCCAACTGATGCGCAACCACTCAATATACCGAGACCAAGC 780
 DB 721 GAGATCAAACTCACAGACCCCCAACTGATGCGCAACCACTCAATATACCGAGACCAAGC 780

QY 781 AGTCATGATGAGGACCTCGCAACATCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
 DB 781 AGTCATGATGAGGACCTCGCAACATCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840

QY 841 ACTTCTGATGCGGTCAACAAAGAGGGCTTTTCATCAACAATGCCACCTCCCTCACCA 900
 DB 841 ACTTCTGATGCGGTCAACAAAGAGGGCTTTTCATCAACAATGCCACCTCCCTCACCA 900

QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAAACAACAACCAATTTCCCAAGATGCTGTGACT 960
 DB 901 CAACCAAGCAGCCACAGCAAGGAGGAAAACAACAACCAATTTCCCAAGATGCTGTGACT 960

QY 961 GAATAGACAAAATAACACAACTGCAACCGTCCATGCCCTCATTAACACTACCAACA 1020
 DB 961 GAATAGACAAAATAACACAACTGCAACCGTCCATGCCCTCATTAACACTACCAACA 1020

QY 1021 ATCTCTACTAACAAACACTCTCCAAACACAACTTCAGCACTCTCTGCAACCAATTACAAAAC 1080
 DB 1021 ATCTCTACTAACAAACACTCTCCAAACACAACTTCAGCACTCTCTGCAACCAATTACAAAAC 1080

QY 1081 ACCAATATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140
 DB 1081 ACCAATATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140

QY 1141 ATACCAACCTGCTCCAAACGGGAATCCCAACACAGCAAGGACGACGAGCAAGAAAA 1200
 DB 1141 ATACCAACCTGCTCCAAACGGGAATCCCAACACAGCAAGGACGACGAGCAAGAAAA 1200

QY 1201 GGGCCGCGCACAAAGCAGCAACCAACAGCAATTCAGCAATTTCCAGTCTCTCCCCACCC 1260
 DB 1201 GGGCCGCGCACAAAGCAGCAACCAACAGCAATTCAGCAATTTCCAGTCTCTCCCCACCC 1260

QY 1261 CCCAGCTGCACTGCAACCAATCTTTGTATATTTTCAAGAAAGGCGATCG 1308
 DB 1261 CCCAGCTGCACTGCAACCAATCTTTGTATATTTTCAAGAAAGGCGATCG 1308

RESULT 3
 ACC57764
 ID ACC57764 standard; cDNA; 2103 BP.
 XX
 AC ACC57764;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Marburg virus envelope glycoprotein coding sequence.
 XX
 KW Lentivirus; pseudotyped virus; gene therapy; vector; gene; ss.
 XX
 OS Marburg virus.
 XX
 FT Key Location/Qualifiers
 FT CDS 10..2055
 FT /+tag= a
 FT /product= "Envelope glycoprotein"
 XX

PN WO2003035849-A2.
 XX 01-MAY-2003.
 PF 28-OCT-2002; 2002WO-US034545.
 PF 26-OCT-2001; 2001US-0353221P.
 PR 26-OCT-2001; 2001US-0356436P.
 XX (MCCR/) MCCRAY P B.
 PA (SAND/) SANDERS D A.
 PA (DAVI/) DAVIDSON B L.
 XX McCray PB, Sanders DA, Davidson BL;
 XX WPI: 2003-421416/39.
 DR P-PSDB; ABR42242.
 XX
 PT New pseudotyped lentivirus comprising a lentiviral capsid, a lipid
 PT bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful
 PT for eliciting an immune response against feline immunodeficiency virus
 PT infection.
 XX
 PS Disclosure; Page 36-39; 4lpp; English.
 XX
 CC The present sequence is the coding sequence for the envelope glycoprotein
 CC of Marburg virus. The invention provides methods for gene transfer to
 CC cells using glycoprotein-pseudotyped lentiviruses. The glycoprotein is
 CC preferably Marburg glycoprotein. The Marburg viral glycoprotein
 CC preferably has a mutation in the C-terminal portion (see ABR42243-47)
 CC that results in a higher titre production of the pseudotyped virus. The
 CC pseudotyped lentivirus is used to introduce nucleic acid sequences
 CC encoding a desired protein into a hepatocyte, brain glial or airway
 CC epithelial cell, in vitro or in vivo. The desired protein is preferably
 CC cystic fibrosis transmembrane conductance regulator, or is the low
 CC density lipoprotein receptor, alpha1-antitrypsin, ornithine
 CC transcarbamylase, Factor VIII or a high affinity glutamate receptor (all
 CC claimed)
 XX
 SQ Sequence 2103 BP; 671 A; 533 C; 408 G; 491 T; 0 U; 0 Other;

Query Match 70.8%; Score 1303.4; DB 8; Length 2103;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTTCCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAATCTCC 60
 DB |||||
 QY 10 ATGAGACACATGTTTCCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAATCTCC 69
 DB |||||
 QY 61 ATTTTAGAGATAGTCTAGTAAATCAACCCCAAAATGTGGATTGGTATGCTCCGGA 120
 DB |||||
 QY 70 ATTTTAGAGATAGTCTAGTAAATCAACCCCAAAATGTGGATTGGTATGCTCCGGA 129
 DB |||||
 QY 121 CTCAGAGACAGAGACGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAATGTT 180
 DB |||||
 QY 130 CTCAGAGACAGAGACGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAATGTT 189
 DB |||||
 QY 181 GATTCCTCTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAT 240
 DB |||||
 QY 190 GATTCCTCTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAT 249
 DB |||||
 QY 241 GTTGAGTACACAGAGGGGAGGAGCAAAAATCATGCTACATATAGTGTAAAGGATCC 300
 DB |||||
 QY 250 GTTGAGTACACAGAGGGGAGGAGCAAAAATCATGCTACATATAGTGTAAAGGATCC 309
 DB |||||
 QY 301 TCTGAAAATCTCTGTGTAGATCTCTCTCAACATCCGTGATCTCCGAAATGCAAA 360
 DB |||||
 QY 310 TCTGAAAATCTCTGTGTAGATCTCTCTCAACATCCGTGATCTCTCAAAATGCAAA 369
 DB |||||
 QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG 420
 DB |||||
 QY 370 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG 429
 DB |||||

QY 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTGTACCGAGGAAAATGTTCTTC 480
 DB |||||
 QY 430 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTGTACCGAGGAAAATGTTCTTC 489
 DB |||||
 QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 540
 DB |||||
 QY 490 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 549
 DB |||||
 QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 600
 DB |||||
 QY 550 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 609
 DB |||||
 QY 601 AGTAAACGGAACGCAAAACGAATGACATGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660
 DB |||||
 QY 610 AGTAAACGGAACGCAAAACGAATGACATGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 669
 DB |||||
 QY 661 ACAAGAACCACCAACATGTGCTCCGTCGCAAAATACCTCCACACTGCCACACGCCGCTCG 720
 DB |||||
 QY 670 ACAAGAACCACCAACATGTGCTCCGTCGCAAAATACCTCCACACTGCCACACGCCGCTCG 729
 DB |||||
 QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCGACGCCAAGC 780
 DB |||||
 QY 730 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCGACGCCAAGC 789
 DB |||||
 QY 781 AGTGATGATGAGGACCTTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
 DB |||||
 QY 790 AGTGATGATGAGGACCTTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 849
 DB |||||
 QY 841 ACTTCTGATGGGTACCAACGCAAGGCTTTTCATCAAAATGCCACCCACTCCCTCACA 900
 DB |||||
 QY 850 ACTTCTGATGGGTACCAACGCAAGGCTTTTCATCAAAATGCCACCCACTCCCTCACA 909
 DB |||||
 QY 901 CAACCAACGACGCCACACAGAGGAGGAAAACAACAAACCAATTCCTCAAGATGCTGTGACT 960
 DB |||||
 QY 910 CAACCAACGACGCCACACAGAGGAGGAAAACAACAAACCAATTCCTCAAGATGCTGTGACT 969
 DB |||||
 QY 961 GAATCTAGACAAAATAAACAACATGACAAACCGTCCATGCCCCCTCATATAACTACCA 1020
 DB |||||
 QY 970 GAATCTAGACAAAATAAACAACATGACAAACCGTCCATGCCCCCTCATATAACTACCA 1029
 DB |||||
 QY 1021 ATCTCTACTACACACCTCCAAAACAACATTCAGCACTCTCTGACCACTTACAAAAC 1080
 DB |||||
 QY 1030 ATCTCTACTACACACCTCCAAAACAACATTCAGCACTCTCTGACCACTTACAAAAC 1089
 DB |||||
 QY 1081 ACCCAATATGACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTCCCTCG 1140
 DB |||||
 QY 1090 ACCCAATATGACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTCCCTCG 1149
 DB |||||
 QY 1141 ATAAACACCTGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACAGAGCAAAA 1200
 DB |||||
 QY 1150 ATAAACACCTGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACAGAGCAAAA 1209
 DB |||||
 QY 1201 GGCCCCGACACACGGGACCAACACGCAAAATGAGCATTTTCACAGTCTCCCTCC 1260
 DB |||||
 QY 1210 GGCCCCGACACACGGGACCAACACGCAAAATGAGCATTTTCACAGTCTCCCTCC 1269
 DB |||||
 QY 1261 CCAGCTGACTGCAACACATCTTGTATATTTTCAAGAAAGCGA 1305
 DB |||||
 QY 1270 CCAGCTGACTGCAACACATCTTGTATATTTTCAAGAAAGCGA 1314
 DB |||||

RESULT 4
 AAD04042
 ID AAD04042 standard; DNA; 2164 BP.
 XX
 AC AAD04042;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE Marburg virus Musoke strain glycoprotein (GP) DNA.
 XX Glycoprotein; GP; immune response; vaccine; antiviral;
 KW type I transmembrane protein; ds.

XX Marburg virus.
OS
XX
FH Key Location/Qualifiers
FT CDS 119..2164
FT FT /*tag= a
FT FT /product= "Marburg virus Musoke strain glycoprotein (GP)"
XX
XX
FN US6200959-B1.
XX
XX 13-MAR-2001.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX (POWDER-) POWDERJECT VACCINES INC.
XX
XX Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX
XX WPI; 2001-280564/29.
DR P-PSDB; AAE00707.
XX
XX Inducing an immune response to a Marburg or Ebola virus involves
PT delivering a genetic vaccine, which contains a genetic construct encoding
PT antigenic determinants for filovirus, using a particle acceleration
PT device.
XX
XX Claim 5; Col 25-30; 33pp; English.
XX
XX The patent discloses a method of inducing an immune response to Marburg
CC or Ebola virus glycoprotein which involves delivering a genetic vaccine,
CC containing a genetic construct encoding antigenic determinants for
CC filovirus, using a particle acceleration device. The genetic vaccine for
CC filovirus is created by joining a DNA sequence encoding at least a
CC portion of the filovirus glycoprotein to a promoter effective to promote
CC transcription of the DNA sequence. This method is useful for inducing
CC humoral, cell-mediated and secretory immune responses in the treated
CC individual. The present sequence is a DNA encoding Marburg virus Musoke
CC strain glycoprotein (GP), which is a type I transmembrane protein. This
CC sequence is used in the construction of genetic vaccine against
XX filoviruses
XX
SQ Sequence 2164 BP; 701 A; 536 C; 412 G; 515 T; 0 U; 0 Other;

Query Match 70.8%; Score 1303.4; DB 4; Length 2164;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGGACAAAATCTCCCC 60
DB 119 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGGACAAAATCTCCCC 178

QY 61 ATTTTAGATAGCTAGTAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAACT 120
DB 179 ATTTTAGATAGCTAGTAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAACT 238

QY 121 CTCGAGACAGAGACGCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 239 CTCGAGACAGAGACGCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 298

QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
DB 299 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 358

QY 241 GTTGTAGTACAGAGGGGAGGAGCCAAAACATGCTACAATTAAGTGTACGGATCCC 300
DB 359 GTTGTAGTACAGAGGGGAGGAGCCAAAACATGCTACAATTAAGTGTACGGATCCC 418

QY 301 TCTGGAATAATCTTGTGTTAGATCTCTCTACCAACATCCGTCATATCCGAAATGCAAA 360
DB 419 TCTGGAATAATCTTGTGTTAGATCTCTCTACCAACATCCGTCATATCCGTCATATCC 478

RESULT 5

ACCT71551
ID ACCT71551 standard; DNA; 7005 BP.

XX ACCT71551;

XX ACCT71551;
XX 10-JUL-2003 (first entry)
XX

QY 361 ACTATCCATCATATTTCAGGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG 420
DB 479 ACTATCCATCATATTTCAGGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG 538

QY 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTATCGGAGGCAAAAGTCTTC 480
DB 539 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTATCGGAGGCAAAAGTCTTC 598

QY 481 ACTGAAGGGAACATAGCAGCTATGATGTCTCAATAGACAGTGCACAAAATGATTTTCTCG 540
DB 599 ACTGAAGGGAACATAGCAGCTATGATGTCTCAATAGACAGTGCACAAAATGATTTTCTCG 658

QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTTAATAATATTGGACAAGT 600
DB 659 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTTAATAATATTGGACAAGT 718

QY 601 AGTAAACGGAACGCAAAACGAATGACACTGATGTTTTCGGCGCTCTTCAAGAATACAAATCT 660
DB 719 AGTAAACGGAACGCAAAACGAATGACACTGATGTTTTCGGCGCTCTTCAAGAATACAAATCT 778

QY 661 ACAAGAAACCAAAACATGTGCTCCGTCGCTCAAAATACCTCCACACTGCCCCACAGCCGTCG 720
DB 779 ACAAGAAACCAAAACATGTGCTCCGTCGCTCAAAATACCTCCACACTGCCCCACAGCCGTCG 838

QY 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAACTCAATACCAAGGACCCCAAGC 780
DB 839 GAGATCAAACTCAAGACACCCCACTGATGCCACCAACTCAATACCAAGGACCCCAAGC 898

QY 781 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 840
DB 899 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 958

QY 841 ACTTCTGATGGGTCCCAAGCAAGGGCTTTCATCAACAATGCGCACCCACTCCCTCACCAC 900
DB 959 ACTTCTGATGGGTCCCAAGCAAGGGCTTTCATCAACAATGCGCACCCACTCCCTCACCAC 1018

QY 901 CAACCAAGCAGCCACAGCAGAGGGAACCAACCAACCATTTCCCAAGATGCTGTGACT 960
DB 1019 CAACCAAGCAGCCACAGCAGAGGGAACCAACCAACCATTTCCCAAGATGCTGTGACT 1078

QY 961 GAACCTAGACAAAATAACACAACTGCACAAACCGTCCATGCCCTCCCTCATAAACAATACCA 1020
DB 1079 GAACCTAGACAAAATAACACAACTGCACAAACCGTCCATGCCCTCCCTCATAAACAATACCA 1138

QY 1021 ATCTCTACTAAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTAACAAAC 1080
DB 1139 ATCTCTACTAAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTAACAAAC 1198

QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCCCT 1140
DB 1199 ACCACCAATGACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCCCT 1258

QY 1141 ATAACAACCTTCCTCCAAACGGAATCCCAACAGCAAGAGAGCAGCAGCAGCAAGCAAAA 1200
DB 1259 ATAACAACCTTCCTCCAAACGGAATCCCAACAGCAAGAGAGCAGCAGCAGCAAGCAAAA 1318

QY 1201 GGCCCCGCCACACAGGCAACCAACACGAAATGAGCAATTTTCCAGTCTCTCCCCCACC 1260
DB 1319 GGCCCCGCCACACAGGCAACCAACACGAAATGAGCAATTTTCCAGTCTCTCCCCCACC 1378

QY 1261 CCAGCTCGACTGCAACAATCTTGTATATTTCAGAAAGAGCGA 1305
DB 1379 CCAGCTCGACTGCAACAATCTTGTATATTTCAGAAAGAGCGA 1423

DE VRC6702 (pVR1012-x/s Marburg GP (dTM)) plasmid.
 XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
 KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
 KW viral infection; filovirus; circular; cyclic; ds.
 XX Synthetic.
 OS WO2003028632-A2.
 XX 10-APR-2003.
 PD 24-SEP-2002; 2002WO-US030251.
 XX 01-OCT-2001; 2001US-0326476P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
 PI WPI; 2003-371961/35.
 DR New bimodal priming and boosting compositions, useful as viral vaccines,
 XX specifically for eliciting an immune response against a filovirus or a
 PT disease caused by infection with filovirus.
 PT Claim 1; Page 188-190; 219pp; English.
 PS The present invention relates to a bimodal priming composition and
 XX boosting composition for priming and boosting an immune response to an
 CC antigen in an individual. The compositions comprise (a) a priming
 CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
 CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
 CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
 CC DNA plasmid selected from AC71521-AC71563 and (b) a boosting
 CC composition comprised of a replication-deficient adenovirus, comprising a
 CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
 CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
 CC epitope-bearing domain, or a replication deficient adenovirus selected
 CC from constructs AC71521-AC71563. The compositions are useful as viral
 CC vaccines, specifically for eliciting an immune response against a
 CC filovirus or a disease caused by infection with filovirus e.g. Ebola
 CC virus infection. This vector expresses the Marburg virus glycoprotein
 CC without its transmembrane and intracellular domains
 XX Sequence 7005 BP; 1931 A; 1758 G; 1588 G; 1728 T; 0 U; 0 Other;
 SQ

Query Match 70.8%; Score 1303.4; DB 10; Length 7005;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTTCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAAATCTCCCC 60
 DB 2034 ATGAGACACATGTTTCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAAATCTCCCC 2093

QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
 DB 2094 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 2153

QY 121 CTCGAGAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
 DB 2154 CTCGAGAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 2213

QY 181 GATTCCTCTTGGGCGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
 DB 2214 GATTCCTCTTGGGCGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 2273

QY 241 GTTGTATACAGAGGGGGAGGAGCAAAAATGCTACAATATAAGTGAACGGATCCC 300
 DB 2274 GTTGTATACAGAGGGGGAGGAGCAAAAATGCTACAATATAAGTGAACGGATCCC 2333

QY 301 TCTGGAAAATCTTGTGTTAGATCTCTTCAACATCCGTAATCCGTAATCCGAAAATGCAAA 360
 DB 301 TCTGGAAAATCTTGTGTTAGATCTCTTCAACATCCGTAATCCGTAATCCGAAAATGCAAA 360

DB 2334 TCTGGAAAATCTTGTGTTAGATCTCTTCAACATCCGTAATCCGTAATCCGAAAATGCAAA 2393

QY 361 ACTATCCATCATATATCAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
 DB 2394 ACTATCCATCATATATCAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 2453

QY 421 GGAGCATTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAAATCTTTC 480
 DB 2454 GGAGCATTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAAATCTTTC 2513

QY 481 ACTGAAGGGACATAGCAGCTATGATTCATTAAGACAGTGCACAAATGATTTTCTCG 540
 DB 2514 ACTGAAGGGACATAGCAGCTATGATTCATTAAGACAGTGCACAAATGATTTTCTCG 2573

QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 600
 DB 2574 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 2633

QY 601 AGTAACGGAAACGCAAAACGAATGACATGGATGTTTGGGGCTCTTTCAAGAATACAAATCT 660
 DB 2634 AGTAACGGAAACGCAAAACGAATGACATGGATGTTTGGGGCTCTTTCAAGAATACAAATCT 2693

QY 661 ACAAGAACCAAAACATGCTCCGTCCTCAAAAATACCTCCACACCTGCCACAGCCCGCTCG 720
 DB 2694 ACAAGAACCAAAACATGCTCCGTCCTCAAAAATACCTCCACACCTGCCACAGCCCGCTCG 2753

QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCAACCAACTCAATACCGAGCCCAAGC 780
 DB 2754 GAGATCAAACTCACAAAGCACCCCAACTGATGCCAACCAACTCAATACCGAGCCCAAGC 2813

QY 781 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
 DB 2814 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 2873

QY 841 ACTTCTGATGGGTCCACCAAGCAAGGGCTTTTCATCAAAATGCCACCCACTCCCTCACA 900
 DB 2874 ACTTCTGATGGGTCCACCAAGCAAGGGCTTTTCATCAAAATGCCACCCACTCCCTCACA 2933

QY 901 CAACCAAGCAGCCACACAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 960
 DB 2934 CAACCAAGCAGCCACACAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 2993

QY 961 GAATCTAGCAAAAATAAACAACATTCGCAACCGTCCATGCCCCCTCATAAACAATACCA 1020
 DB 2994 GAATCTAGCAAAAATAAACAACATTCGCAACCGTCCATGCCCCCTCATAAACAATACCA 3053

QY 1021 ATCTCTACTTACCAACCTCCAAACACAACTTCAGCACTCTCTGCAACCATTTACAAAC 1080
 DB 3054 ATCTCTACTTACCAACCTCCAAACACAACTTCAGCACTCTCTGCAACCATTTACAAAC 3113

QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 1140
 DB 3114 ACCACCAATGACACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 3173

QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGAGCAAAAAA 1200
 DB 3174 ATAAACAACCTGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGAGCAAAAAA 3233

QY 1201 GGCCCCGCCACACAGGACCAACACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCCACC 1260
 DB 3234 GGCCCCGCCACACAGGACCAACACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCCACC 3293

QY 1261 CCCAGCTCGATGCAACACATCTTGTATATTTTCAAGAAGAAAGCGA 1305
 DB 3294 CCCAGCTCGATGCAACACATCTTGTATATTTTCAAGAAGAAAGCGA 3338

RESULT 6
 ACC71550
 ID ACC71550 standard; DNA; 7778 BP.
 XX AC ACC71550;
 XX

DT 10-JUL-2003 (first entry)
DE VRC6701 (pVR1012-Marburg) plasmid.
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX Synthetic.
OS
XX
XX WO2003028632-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
XX WPI; 2003-371961/35.
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 185-187; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, or
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence has a Marburg glycoprotein (GP)
CC open reading frame, Musoke strain, cloned into a VRC6700 backbone
XX
XX Sequence 7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;

Query Match 70.8%; Score 1303.4; DB 10; Length 7778;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTCTCTATCAGTCTTATCTTAATTCAAGGGACAAAAATCTCCCC 60
DB 2034 ATGAAGACACATGTTCTCTATCAGTCTTATCTTAATTCAAGGGACAAAAATCTCCCC 2093

QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAACT 120
DB 2094 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAACT 2153

QY 121 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTTCACATGAGTGGGCAAAAAGTTGCT 180
DB 2154 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTTCACATGAGTGGGCAAAAAGTTGCT 2213

QY 181 GATTTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
DB 2214 GATTTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 2273

QY 241 GTTGAAGTACAGAGGGGGAGGAGCAAAAACATGCTACAATATAGTAAAGTAAAGGATCCC 300
DB 2274 GTTGAAGTACAGAGGGGGAGGAGCAAAAACATGCTACAATATAGTAAAGTAAAGGATCCC 2333

301 TCTGGAATACTCTTGTCTAGATCCTCTACCAACATCCGTGACTATCCGAATGCAAA 360
2334 TCTGGAATACTCTTGTCTAGATCCTCTACCAACATCCGTGACTATCCGAATGCAAA 2393

QY 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCAAGGGGATCGCCCTTCAATTAATGG 420
DB 2394 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCAAGGGGATCGCCCTTCAATTAATGG 2453

QY 421 GGAGCATTTTCTGTATGATCGCATTTGCTCTCCCAACAATGTATCCGAGGCAAAAGTCTTC 480
DB 2454 GGAGCATTTTCTGTATGATCGCATTTGCTCTCCCAACAATGTATCCGAGGCAAAAGTCTTC 2513

QY 481 ACTGAAGGGACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTTCG 540
DB 2514 ACTGAAGGGACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTTCG 2573

QY 541 CGGCAAGGACAAAGGTACCGTCAATATGAACTCTGACTTCTACTATAAATATTTGGACAAGT 600
DB 2574 CGGCAAGGACAAAGGTACCGTCAATATGAACTCTGACTTCTACTATAAATATTTGGACAAGT 2633

QY 601 AGTAAACGGAACGCAAAACGAATGACATGGATGTTTTCGGCGCTCTTCAAGAATACAAATTTCT 660
DB 2634 AGTAAACGGAACGCAAAACGAATGACATGGATGTTTTCGGCGCTCTTCAAGAATACAAATTTCT 2693

QY 661 ACAAGAACCAAAACATGTGCTCCGTCCTCAAAATATCTCCACACTGCCCCACAGCCCGTCGG 720
DB 2694 ACAAGAACCAAAACATGTGCTCCGTCCTCAAAATATCTCCACACTGCCCCACAGCCCGTCGG 2753

QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATATACCAGGACCAAGC 780
DB 2754 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATATACCAGGACCAAGC 2813

QY 781 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 840
DB 2814 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 2873

QY 841 ACTTCTGATGCGGTTCACCAAGCAGGGGCTTTCATCAACATGCCACCCACTCCCTCACA 900
DB 2874 ACTTCTGATGCGGTTCACCAAGCAGGGGCTTTCATCAACATGCCACCCACTCCCTCACA 2933

QY 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
DB 2934 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 2993

QY 961 GAATAGACAAAAATAACACAACATTCGCAACACCGTCCATGCCCCCTCATPAACACTACACA 1020
DB 2994 GAATAGACAAAAATAACACAACATTCGCAACACCGTCCATGCCCCCTCATPAACACTACACA 3053

QY 1021 ATCTCTACTAACAACACCTCCAAAACAACTTACAGCACTCTCTGCAACCAATTCACAAAC 1080
DB 3054 ATCTCTACTAACAACACCTCCAAAACAACTTACAGCACTCTCTGCAACCAATTCACAAAC 3113

QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCCCCTCG 1140
DB 3114 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCCCCTCG 3173

QY 1141 ATAAACAACCTTCCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACCAGCAGCAAAAAA 1200
DB 3174 ATAAACAACCTTCCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACCAGCAGCAAAAAA 3233

QY 1201 GGCCCCGGCACAACGGGACCAAAACACGCAAAATGAGCAATTCACAGTCTCTCCCCCACC 1260
DB 3234 GGCCCCGGCACAACGGGACCAAAACACGCAAAATGAGCAATTCACAGTCTCTCCCCCACC 3293

QY 1261 CCCAGCTCGACTGCAACAACATCTTTGTATATTTTCAGAGAAAGCGA 1305
DB 3294 CCCAGCTCGACTGCAACAACATCTTTGTATATTTTCAGAGAAAGCGA 3338

RESULT 7
ID ACC71552
XX

```
AC ACCT1552;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6710 (pAdapt Marburg GP(dtm)) plasmid.
XX
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
XX WO2003028632-A2.
PN
XX
XX 10-APR-2003.
PD
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI
XX
XX WPI; 2003-371961/35.
DR
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 190-193; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence is an adenovirus shuttle vector,
CC expressing the Marburg virus glycoprotein without its transmembrane and
CC intracellular domains
XX
XX Sequence 8256 BP; 2048 A; 2049 C; 2108 G; 2051 T; 0 U; 0 Other;
SQ
Query Match 70.8%; Score 1303.4; DB 10; Length 8256;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGGACAAAAATCTCCCC 60
DB 1431 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGGACAAAAATCTCCCC 1490
QY 61 ATTTAGAGATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT 120
DB 1491 ATTTAGAGATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT 1550
QY 121 CTCCAGACAGACAGACCTCCATCTGTGGATTCACACTGAGTGGGCAAAAGTGTCT 180
DB 1551 CTCCAGACAGACAGACCTCCATCTGTGGATTCACACTGAGTGGGCAAAAGTGTCT 1610
QY 181 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT 240
DB 1611 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT 1670
QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAAAATGCTACATATAAGTGTAAACGGATCCC 300
```

RESULT 8

AAZ87211
ID AAZ87211 standard; CDNA; 11460 BP.
XX
AC AAZ87211;
XX
15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX
DE VEE replicon comprising Marburg virus genes.
XX
KW VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV; filovirus; glycoprotein; GP; nucleoprotein; NP; structural protein; VP35; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX
OS Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.
OS Chimeric.
XX
FH Location/Qualifiers
CDS 104..2182
FT /*tag= a
FT /product= "Marburg virus nucleoprotein (NP, AAY77128)"
FT CDS 2944..3933
FT /*tag= b
FT /product= "Marburg virus structural protein VP35 (AAY77130)"
FT CDS 4567..5478
FT /*tag= c
FT /product= "Marburg virus structural protein VP40 (AAY77129)"
FT CDS 5940..7985
FT /*tag= d
FT /product= "Marburg virus glycoprotein (GP, AAY77127)"
FT CDS 8864..9697
FT /*tag= e
FT /product= "Marburg virus structural protein VP30 (AAY77131)"
FT CDS 10200..10961
FT /*tag= f
FT /product= "Marburg virus structural protein VP24 (AAY77132)"
XX
PN WO200000616-A2.
XX
PD 06-JAN-2000.
XX
PF 21-JUN-1999; 99WO-US014174.
XX
PR 29-JUN-1998; 98US-0091403P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;
XX
DR WPI; 2000-160676/14.
DR P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132, AAY77133.
XX
XX Novel Marburg virus vaccines used to induce an immune response against the infection in nonhuman primates.
XX
PS Claim 1; Page 56-57; 57pp; English.
XX
XX The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine of the invention comprises a Venezuelan equine encephalitis virus (VEE) replicon containing copies of the genes encoding Marburg virus (Musoke strain) glycoprotein (GP), a deletion mutant GP (GP-Delta-TM), nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24. The Marburg virus genes replace the VEE virus structural protein genes; the result is a self-replicating RNA molecule that encodes its own replicase and transcriptase functions, and in addition makes abundant quantities of the Marburg virus proteins. When replicon RNA is

transfected into eukaryotic cells, along with two helper RNAs that express the VEE virus structural proteins, the replicon RNA is packaged into VEE virus-like particles by the VEE virus structural proteins, which are provided in trans. Since the helper RNAs lack packaging signals necessary for further propagation, the resulting VEE replicon particles (VRPs) which are produced are infectious for one cycle but are defective thereafter. On VRP infection of a cell, an abortive infection occurs whereby the cell produces the Marburg virus proteins, is ultimately killed by the infection, but does not produce any viral progeny. The VEE replicon provides a potent tool for vaccination with Marburg virus antigens. The replicons, vectors and constructs are used to produce vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic fever) in mammals, to elicit immune responses against Marburg antigens, to confer protective immunity, and to reduce disease symptoms and reduce the severity of disease. Studies of non-human primates vaccinated with the replicon indicate that the vaccine will be efficient in protecting humans against Marburg virus. Prior art Marburg virus vaccination strategies have used formalin-inactivated Marburg virus, which are only partially successful at protecting against Marburg virus infection. The present sequence represents a cDNA corresponding to a VEE virus replicon encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24.
(Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;
Query Match 70.8%; Score 1303.4; DB 3; Length 11460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC 60
DB 5940 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC 5999
QY 61 ATTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGTCCTCGGAAT 120
DB 6000 ATTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGTCCTCGGAAT 6059
QY 121 CTCAGAAAGACAGAAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCAGAAAGACAGAAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGGTGTACCTCCCAAGAAAT 240
DB 6120 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGGTGTACCTCCCAAGAAAT 6179
QY 241 GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACATATAAGTGTAAAGGATCCC 300
DB 6180 GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACATATAAGTGTAAAGGATCCC 6239
QY 301 TCTGAAAAATCCTTGCTGTTAGATCCTCTTACCAACATCCTCGTACTATCCGAAATGCAAA 360
DB 6240 TCTGAAAAATCCTTGCTGTTAGATCCTCTTACCAACATCCTCGTACTATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
DB 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAAGGGAACATAGCAGCTTATGATTGTCAATAGACAGTGCACAAAAATGATTTCTCG 540
DB 6420 ACTGAAGGGAACATAGCAGCTTATGATTGTCAATAGACAGTGCACAAAAATGATTTCTCG 6479
QY 541 CGGCAGGACAGAGGTACCGTCATATGATCTGACTTCTACTTAATAATATTTGGCAAGT 600
DB 6480 CGGCAGGACAGAGGTACCGTCATATGATCTGACTTCTACTTAATAATATTTGGCAAGT 6539
QY 601 AGTAACGGAACGCAAAACGAATGACACTGGATGTTTCCGGCGCTCTTCAAGAATACAAATCT 660
DB 6540 AGTAACGGAACGCAAAACGAATGACACTGGATGTTTCCGGCGCTCTTCAAGAATACAAATCT 6599

```

QY 661 ACAAGAACCAAAATGCTCGTCCAAAATACTCCACCACTGCCACAGCCCGTCCG 720
Db 6600 ACAAGAACCAAAATGCTCGTCCAAAATACTCCACCACTGCCACAGCCCGTCCG 6659
QY 721 GAGATCAAACTCACAGACCCCACTGATGCGGAGGAGGAGGAGGAGGAGGAGG 780
Db 6660 GAGATCAAACTCACAGACCCCACTGATGCGGAGGAGGAGGAGGAGGAGGAGG 6719
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGGAGGAGGAGGAGG 840
Db 6720 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGGAGGAGGAGGAGG 6779
QY 841 ACTTCTGATGCGGTCAACGAAGGCTTTTATCAACAATGCCACCCACTCCCTCACA 900
Db 6780 ACTTCTGATGCGGTCAACGAAGGCTTTTATCAACAATGCCACCCACTCCCTCACA 6839
QY 901 CAACCAAGCACCCACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 6840 CAACCAAGCACCCACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6899
QY 961 GAACTAGACAAAATAAACAACCTGACAAACCGTCCATGCCCTCATAACTACCACA 1020
Db 6900 GAACTAGACAAAATAAACAACCTGACAAACCGTCCATGCCCTCATAACTACCACA 6959
QY 1021 ATCTCTACTAACAACACTCCAAACAACTTTGACACTCTCTCTGACCACTTACAAAC 1080
Db 6960 ATCTCTACTAACAACACTCCAAACAACTTTGACACTCTCTCTGACCACTTACAAAC 7019
QY 1081 ACCACCAATGACACACACAGAGCAATCACTGAAATGAGCAAAACCGTCCCTCG 1140
Db 7020 ACCACCAATGACACACACAGAGCAATCACTGAAATGAGCAAAACCGTCCCTCG 7079
QY 1141 ATAACAACCTGCTCTCAACGGGAAATCCACACAGCAAGAGCAGCAGCAAGGAG 1200
Db 7080 ATAACAACCTGCTCTCAACGGGAAATCCACACAGCAAGAGCAGCAGCAAGGAG 7139
QY 1201 GGCCCCGGCACAACGGCAGCAACAGCAAGCAAAATGAGCAATTTCCAGGTCCTCC 1260
Db 7140 GGCCCCGGCACAACGGCAGCAACAGCAAGCAAAATGAGCAATTTCCAGGTCCTCC 7199
QY 1261 CCAGCTGCACTGCACACACATCTTGTATTTTTCAGNAGAAAGCGA 1305
Db 7200 CCAGCTGCACTGCACACACATCTTGTATTTTTCAGNAGAAAGCGA 7244

```

RESULT 9

AZ51039
ID AZ51039 standard; DNA; 2100 BP.

AC AZ51039;

DT 05-JUN-2000 (first entry)

DE Marburg virus envelope glycoprotein DNA.

KW Marburg virus envelope glycoprotein; viral glycoprotein;
KW pseudotyped retrovirus; MLV; Moloney murine leukemia virus; antibody;
KW transduction; screening agent; immunological agent;
KW pharmacological agent; ss.

OS Marburg virus.

EH Key Location/Qualifiers

FT CDS 10..2052

FT /*tag= a

FT /product= "Marburg virus envelope glycoprotein"

XX W0200008131-A2.

PN 17-FEB-2000.

PD 04-AUG-1999; 99WO-US017702.

XX

PR 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
XX
PA (PURD) PURDUE RES FOUND.
XX
PI Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
PI Fischbach MA;
XX
XX WPI; 2000-224030/19.
XX P-PSDB; AAV70075.
XX
XX Cells that produce inventive pseudotyped retroviruses having a broad host
XX range useful for introducing nucleotide sequences into target cells.
XX
XX Example 11; Page 60-61; 65pp; English.
XX
XX The present sequence encodes the Marburg virus envelope glycoprotein.
XX Plasmid pMBGFI was produced by cloning into the plasmid pSP72 nucleotide
XX sequences corresponding to nucleotides 5931-8033 from the Marburg virus
XX genome. This is used for generation of stable cell lines transiently
XX producing Marburg-MLV (Moloney murine leukemia virus) pseudotyped
XX retrovirus with a broad host range. These cells having different viral
XX glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
XX transduced with a desired ribonucleotide sequence. This is used to
XX identify screening agents effective in blocking viral entry into a cell.
XX These agents may be immunological agents like monoclonal or polyclonal
XX antibodies. The pharmacological agents include proteins, peptides or
XX various chemical agents. The pseudotyped retrovirus may be useful in
XX methods of identifying cell surface receptors that allow viral entry
XX
SQ Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;

Query Match 69.9%; Score 1287.4; DB 3; Length 2100;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAAGACCAACATGTTTCTTATCAGTCTTATCTTTAATTCAGGAGCAAAAAATCTCCC 60

Db 10 ATGAAGACCAACATGTTTCTTATCAGTCTTATCTTTAATTCAGGAGCAAAAAATCTCCC 69

QY 61 ATTTTAGAGTAGTAGTAAATTAATCAACCCCAAAATGCGATTCTCGTATCTCCGAACT 120

Db 70 ATTTTAGAGTAGTAGTAAATTAATCAACCCCAAAATGCGATTCTCGTATCTCCGAACT 129

QY 121 CTCAGAGACAGAGACGTCCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTCT 180

Db 130 CTCAGAGACAGAGACGTCCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTCT 189

QY 181 GATTCCCTTTGGAGGCAATCCAAGCGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAT 240

Db 190 GATTCCCTTTGGAGGCAATCCAAGCGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAT 249

QY 241 GTTGAGTACAGAGGGGGAGGAGCCAAACATCTACATATAGTGTACAGGATCC 300

Db 250 GTTGAGTACAGAGGGGGAGGAGCCAAACATCTACATATAGTGTACAGGATCC 309

QY 301 TCTGAAAAATCTCTGTTAGATCTCTTACCAACATCCGTCAGTATCCGAAATGCAAA 360

Db 310 TCTGAAAAATCTCTGTTAGATCTCTTACCAACATCCGTCAGTATCCGAAATGCAAA 369

QY 361 ACTATCCATCATATTCAAAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 420

Db 370 ACTATCCATCATATTCAAAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 429

QY 421 GGAGCATTTTTCTGTATGATGCGATTGCTCCCAACAATGTACGGGAGGAACTTC 480

Db 430 GGAGCATTTTTCTGTATGATGCGATTGCTCCCAACAATGTACGGGAGGAACTTC 489

QY 481 ACTGAAGGGCAACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540

Db 490 ---GAAGGGACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTCTCG 546

QY 541 CGGCAAGGACAGGGTACCGTCATATGAATCTGACTTCTTAATAAATATTTGGACAAGT 600

```
Db 547 CGGCAAGGACAGGGTACCGTCTATGATCTGACTTCTACTATAATATTTGGCAAGT 606
Qy 601 AGTAAACGGAAACGAAACGAATGACACTGGATGTTTTCGGCGCTTTCAAGAATATCAATTTCT 660
Db 607 AGTAAACGGAAACGAAACGAATGACACTGGATGTTTTCGGCGCTTTCAAGAATATCAATTTCT 666
Qy 661 ACARAGAACCAACATGCTGCTCGTCCAAATATCTCCACCACTGCCCCACAGCCCGTCCG 720
Db 667 ACARAGAACCAACATGCTGCTCGTCCAAATATCTCCACCACTGCCCCACAGCCCGTCCG 726
Qy 721 GAGATCAAACTCACAAGACACCCCAACTGATGCGCACCAAACTCAATACCAAGACCAAGC 780
Db 727 GAGATCAAACTCACAAGACACCCCAACTGATGCGCACCAAACTCAATACCAAGACCAAGC 786
Qy 781 AGTGATGATGAGGACCTCGCAACATCCCGCTCAGGCTCGGAGAACGAGAACCCCAACACA 840
Db 787 AGTGATGATGAGGACCTCGCAACATCCCGCTCAGGCTCGGAGAACGAGAACCCCAACACA 846
Qy 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAAAATGCGCACTCCCTCCCTACCA 900
Db 847 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAAAATGCGCACTCCCTCCCTACCA 906
Qy 901 CAACCAAGCACGCCACAGCAAGGAGGAAACACACAAACCATTTCCTCAAGATGCTGTGACT 960
Db 907 CAACCAAGCACGCCACAGCAAGGAGGAAACACACAAACCATTTCCTCAAGATGCTGTGACT 966
Qy 961 GAACTAGACAAATAAACAACACTGCGACACCGTCCATGCCCGCTTCATAAACAACCTACCA 1020
Db 967 GAACTAGACAAATAAACAACACTGCGACACCGTCCATGCCCGCTTCATAAACAACCTACCA 1026
Qy 1021 ATCTCTACTAACAACACCTCCAAACACAACTTCAGCACTCTCTGCGACCAATACAAAC 1080
Db 1027 ATCTCTACTAACAACACCTCCAAACACAACTTCAGCACTCTCTGCGACCAATACAAAC 1086
Qy 1081 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAAAACAGTGCCCGCTCG 1140
Db 1087 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAAAACAGTGCCCGCTCG 1146
Qy 1141 ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGAGCAAGCAAAAA 1200
Db 1147 ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGAGCAAGCAAAAA 1206
Qy 1201 GGGCCCGCCCAACAGGACCAAAACAGCAAAATGAGCAATTTACAGTCTCCCGCCAC 1260
Db 1207 GGGCCCGCCCAACAGGACCAAAACAGCAAAATGAGCAATTTACAGTCTCCCGCCAC 1266
Qy 1261 CCCAGCTCCAGTGCACAACTCTGTATATTTTCAGAGAAAGCGA 1305
Db 1267 CCCAGCTCCAGTGCACAACTCTGTATATTTTCAGAGAAAGCGA 1311
```

RESULT 10

ABT13454
ID ABT13454 standard; DNA; 2051 BP.

XX AC ABT13454;

XX DT 30-JAN-2003 (first entry)

DE DNA encoding a chimeric filovirus protein MBGV-GP1/GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

XX Immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

OS Chimeric.

XX WO200279239-A2.

XX 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.
XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABJ18476.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Disclosure; Page 81-84; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
XX Sequence 2051 BP; 655 A; 520 C; 400 G; 476 T; 0 U; 0 Other;

Query Match 69.2%; Score 1274; DB 10; Length 2051;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 1296; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy 1 ATGAAGACCACTGTTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAAATCTCCCC 60
Db 10 ATGAAGACCACTGTTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAAATCTCCCC 69
Qy 61 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGTCGGAAC 120
Db 70 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGTCGGAAC 129
Qy 121 CTCAGAGACAGAAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGTC 180
Db 130 CTCAGAGACAGAAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGTC 189
Qy 181 GATTCCTCTTGGAGGATCCAGCGATGGCTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 190 GATTCCTCTTGGAGGATCCAGCGATGGCTTTCAGGACAGGTGTACCTCCCAAGAT 249
Qy 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACAATATAAGTAAACCGATCCC 300
Db 250 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACAATATAAGTAAACCGATCCC 309
Qy 301 TCTGCAAAATCTTGTGTTAGATCTCTCTACCAACATCCGTGACTATCCGAATGCAA 360
Db 310 TCTGCAAAATCTTGTGTTAGATCTCTCTACCAACATCCGTGACTATCCGAATGCAA 369
Qy 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGATCGCCCTTATTATGG 420
Db 370 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGATCGCCCTTATTATGG 429
Qy 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCCAACATGTAACCGAGGCAAAAGTCTTC 480
Db 430 GGAGCATTTTTCTGTATGATCGCATTT-----ATGTACCGAGGCAAAAGTCTTC 477
Qy 481 ACTGAAGGAAACATAGCAGCTATGATGTCATTAAGACAGTGCACAAATGATTTCTCG 540
Db 478 ACTGAAGGAAACATAGCAGCTATGATGTCATTAAGACAGTGCACAAATGATTTCTCG 537
Qy 541 CGGCAAGGACAGGGTACCGTCAATGAACTGCTTCTACTATAATATATGCAACAGT 600
Db 538 CGGCAAGGACAGGGTACCGTCAATGAACTGCTTCTACTATAATATATGCAACAGT 597
Qy 601 AGTAAACGCAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAA 660
Db 598 AGTAAACGCAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAA 657


```
QY 661 ACAAGAACCAACATGTGCTCGGTCCAAATACTCTCCACCACTGCGCCACAGCCCGTCCG 720
D 658 ACAAGAACCAACATGTGCTCGGTCCAAATACTCTCCACCACTGCGCCACAGCCCGTCCG 717
QY 721 GAGATCAAACTCAAGACACCCCACTGATGCCAACAACTCAATACACGACCCAGC 780
D 718 GAGATCAAACTCAAGACACCCCACTGATGCCAACAACTCAATACACGACCCAGC 777
QY 781 AGTGATGATGAGACCTCGCAACATCGGCTCAGGGTCGGAGAGAGGAAACCCACACA 840
D 778 AGTGATGATGAGACCTCGCAACATCGGCTCAGGGTCGGAGAGAGGAAACCCACACA 837
QY 841 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACC 900
D 838 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACC 897
QY 901 CAACCAAGCAGCCACAGCAAGAGAGGAAACACACAAACCAATCCCAAGATCTGTGACT 960
D 898 CAACCAAGCAGCCACAGCAAGAGAGGAAACACACAAACCAATCCCAAGATCTGTGACT 957
QY 961 GAACTAGACAAAATAACACAACCTGCAACACCGTCCATGCCCTCATCAACTACTACCACA 1020
D 958 GAACTAGACAAAATAACACAACCTGCAACACCGTCCATGCCCTCATCAACTACTACCACA 1017
QY 1021 ATCTCTACTAACAACACCTCCAAACACAACACTTTCAGCACTCTCTGCGACCATTCACAAAC 1080
D 1018 ATCTCTACTAACAACACCTCCAAACACAACACTTTCAGCACTCTCTGCGACCATTCACAAAC 1077
QY 1081 ACACAATGACACACACAGAGCAACATCACTGAAATGAGCAACACAGTCCGCCCTCG 1140
D 1078 ACCACAATGACACACACAGAGCAACATCACTGAAATGAGCAACACAGTCCGCCCTCG 1137
QY 1141 ATAACAACCTGCTCCCAACGGGAATCCACACAGCAAGAGCAGCAGCAGCAAAA 1200
D 1138 ATAACAACCTGCTCCCAACGGGAATCCACACAGCAAGAGCAGCAGCAGCAAAA 1197
QY 1201 GGCCCGCCACACGCGCACCAACAGCAACAAATGAGCATTTTCCAGCATCTCTCCGCCACC 1260
D 1198 GGCCCGCCACACGCGCACCAACAGCAACAAATGAGCATTTTCCAGCATCTCTCCGCCACC 1257
QY 1261 CCAGTCTGACTGACACAACTCTGTATATTTTCAGAGAAAGGATCG 1308
D 1258 CCAGTCTGACTGACACAACTCTGTATATTTTCAGAGAAAGGATCG 1305

RESULT 11
ADM48345
ID ADM48345 standard; DNA; 19112 BP.
AC ADM48345;
XX
XX
DT 18-NOV-2004 (first entry)
XX
XX Marburg virus viral protein genomic DNA.
DE
XX
XX Filovirus; viral protein; NP protein; VP30 protein; VP35 protein;
KW VP40 protein; L protein; membrane-associated protein; matrix protein;
KW polymerase complex protein; minor nucleoprotein; glycoprotein; GP;
KW viral mutagenesis study; vaccine; gene therapy; gene; ds.
XX
OS Marburg virus.
XX
XX Key Location/Qualifiers
PH 103..2190
FT CDS /*tag= a
FT /*product= "NP protein"
FT 2944..3933
FT CDS /*tag= b
FT /*product= "VP35 protein"
FT 4567..5478
FT CDS /*tag= c
FT /*product= "VP40 protein"
```

```
FT CDS 5940..7985
FT /*tag= d
FT /*product= "GP protein"
FT 8867..9712
FT CDS /*tag= e
FT /*product= "VP30 protein"
FT 10205..10966
FT CDS /*tag= f
FT /*product= "VP24 protein"
FT 11479..18474
FT CDS /*tag= g
FT /*product= "L protein"
XX
XX US2003215794-A1.
PN 20-NOV-2003.
PD
XX
XX 29-JAN-2003; 2003US-00353856.
XX
XX 31-JAN-2002; 2002US-0353972P.
XX
XX (KAWA/) KAWAOKA Y.
PA (JASE/) JASENOSKY L D.
PA (NEUM/) NEUMANN G.
XX
XX Kawaoka Y, Jasenosky LD, Neumann G;
PI WPI; 2004-010776/01.
XX
XX P-PSDB; ADM48338, ADM48339, ADM48340, ADM48341, ADM48342, ADM48343,
DR ADM48344.
XX
XX Preparing filovirus useful in producing vaccines and gene therapy
PT vectors, comprises contacting a cell with a vector comprising a promoter
PT operably linked to a filovirus genomic cDNA or protein.
XX
XX Disclosure; SEQ ID NO 27; 104pp; English.
XX
XX The invention relates to a method of preparing filovirus. The method
XX involves contacting a cell with a vector comprising a promoter operably
XX linked to a filovirus genomic cDNA or its portion linked to a
XX transcription termination sequence or to a DNA segment encoding a
XX filovirus RNA transcriptase-polymerase, filovirus NP, VP30 or VP35, to
XX yield infectious filovirus. The portion of the cDNA when transcribed
XX yields a RNA capable of being packaged into filovirus virions or capable
XX of being replicated in the presence of filovirus proteins. The method is
XX useful for producing filovirus useful in viral mutagenesis studies and in
XX the production of vaccines and gene therapy vectors. The present sequence
XX is marburg virus (NP; L; polymerase complex-VP35; matrix protein-VP40;
XX glycoprotein-GP; minor ncleoprotein-VP30; membrane-associated protein-
XX VP24) genomic DNA.
XX
XX SQ Sequence 19112 BP; 6176 A; 3757 C; 3578 G; 5601 T; 0 U; 0 Other;
Query Match 62.4%; Score 1148.2; DB 13; Length 19112;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTTCTTATCAGTCTTATTAATTAAGGACAAAAAATCTCC 60
D 5940 ATGAAGACCACATGTTTCTTATCAGTCTTATTAATTAAGGACAAAAAATCTCC 5999
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTC 120
D 6000 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTC 6059
QY 121 CTCGAGAGACAGAGACGTCCATCTGATGGATTCACACTGAGTGGGCAAAAGTTC 180
D 6060 CTCGAGAGACAGAGATGTCCATCTGATGGATTCACACTGAGTGGGCAAAAGTTC 6119
QY 181 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAT 240
D 6120 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAT 6179
```



```
QY 241 GTTGTAGTACAGAGGGGAGGAGCCAAACATCTCTACATATATAGTGTACGGATCCC 300
Db 6180 GTTGTAGTATACAGAGGGGAGGAGCCAAACATCTCTACATATATAGTGTACGGATCCC 6239
QY 301 TCTGGAATAATCTGTCTGTAGATCCTCTCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 6240 TCTGGAATAATCTGTCTGTAGATCCTCTCAACATCCGCTGACTATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGTCACAGGGGATCGCCCTTCATTATGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGTCGCAAGGGATCGCCCTTCATTATGG 6359
QY 421 GGAGCATTTTTCTGTATGATCGATTCGCTCCACAAATGTACCGAGGCAAGTCTTC 480
Db 6360 GGAGCATTTTTCTGTATGATCGATTCGCTCCACAAATGTACCGAGGCAAGTCTTC 6419
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGATGCACAAAATGATTTCTCG 540
Db 6420 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGATGCACAAAATGATTTCTCG 6479
QY 541 CGGCAAGGACAAAGGTACCGTCATATGATCTGACTTCTTACTTAATAATATTTGGCAAGT 600
Db 6480 AGGCAAGGACAAAGGTACCGTCATATGATCTGACTTCTTACTTAATAATATTTGGCAAGT 6539
QY 601 AGTAACGGNAACGCAACGNAATGACACTGGATGTTTCGGGGCTCTTCAAGAAATACAAATCT 660
Db 6540 AACTAATGGAACACAAACGNAATGACACTGGATGTTTCGGGGCTCTTCAAGAAATACAAATCT 6599
QY 661 ACAAGAGAACCAACATGTCTCGTCCAAATACCTCCACCACTCCGACAGCCCGCTCG 720
Db 6600 ACGAAGATCAAAACATGTCTCGTCCAAATACCTCCACCACTCCGACAGCCCGCTCG 6659
QY 721 GAGATCAAACTCAAGCACCCCAACTGATGTCACCAAACTCAATPACACGAGCCCAAGC 780
Db 6660 GAGATCAAACTCAAGCACCCCAACTGATGTCACCAAACTCAATPACACGAGCCCAAGC 6719
QY 781 AGTGATGATGAGGACTGTGCAACATCGGCTCAGGGTCGGAGAAAGAGAACCCACACA 840
Db 6720 AATGATGATGAGGACTGTGCAACATCGGCTCAGGGTCGGAGAAAGAGAACCCCTTATACA 6779
QY 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAAAATGCCACCACTCCCTCACC 900
Db 6780 ACTTCTGATGCGGTCACTAAGCAAGGGCTTTTCATCAAAATGCCACCACTCCCTCACC 6839
QY 901 CAACCAAGCACGCCACAGCAGGAGGAAACACACAAACCAATTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCACGCCACAGCAGGAGGAAACACACAGACCAATTTCCCAAGGTAAGTGTGACT 6899
QY 961 GAACTAGACAAAATAACACAACTGACAAACCGTCCATGCGCCCTCATATAACACTACACA 1020
Db 6900 GAAACCAACAAACCAACACAAACGCGCACAAACCGTCCATGCGCCCTCATATAACACTACACA 6959
QY 1021 ATCTCTACTTAAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACCAAAAC 1080
Db 6960 ATCTCTACTTAAACACACCTCCAAAGAACAACTTCAGCACTCTCTCTGTAATCACTACAAAC 7019
QY 1081 ACCCAATATGACACACAGAGCAAACTCACTGAAATATGACAAACAGTGCCTCCCTCG 1140
Db 7020 ACCCAATATGACACACAGAGCAAGCACTGAAATATGACAAACAGTGCCTCCCTCG 7079
QY 1141 ATAAACCTCTGCTCCAAACGGGAAATCCACACAGCAAAAGAGACCAAGAGCAAGGAAAA 1200
Db 7080 AAAACAACTCTGCTCCAAACAGGAAATCTTACCACAGCAAAAGAGCACTTAAACACAGAAA 7139
QY 1201 GGGCCCGCCCAACCGGACCAACAGCAAAATGAGCAATTCACAGTCTCTCCCTCCCTCC 1260
Db 7140 GGGCCCGCCCAACCGGACCAACAAATATGACAAATGGGCAATTTAAACAGTCTCTCCCTCCCTCC 7199
QY 1261 CCAGCTGCACTGCAACCACTCTGTATATTTTCAAGAAAGCGGA 1305
Db 7200 CCCAACCCGACCAACCACTCTGTATATTTTCAAGAAAGACGA 7244
```

```
RESULT 12
ABT13453
ID ABT13453 standard; DNA; 2046 BP.
XX AC ABT13453;
XX DT 30-JAN-2003 (first entry)
XX DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX Unidentified.
XX OS Chimeric.
XX PN WO200279239-A2.
XX PD 10-OCT-2002.
XX PF 31-JAN-2002; 2002WO-US003339.
XX PR 31-JAN-2001; 2001US-0267522P.
XX PA (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX PI Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABJ18475.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Claim 18; Page 78-80; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
SQ Sequence 2046 BP; 662 A; 486 C; 379 G; 519 T; 0 U; 0 Other;

Query Match 42.4%; Score 780; DB 10; Length 2046;
Best Local Similarity 74.8%; Pred. No. 5.2e-223;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATCTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCATATATTTTCTGATGTCTCATTTTATTCAAAGTATAAAAATCTCTCCT 60
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAATTTGTAGTAACAGCAACCTCAAGATGTAGATTCACTGTGCTCCGGAAC 120
QY 121 CTCAGAGACAGAGAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTTGCT 180
Db 121 CTCGAAAGACAGAGAAGATGTTTCACTGTGGGATTTACCTGAGTGGGCAAAAGTTGCT 180
QY 181 GATTTCCTTTTGGAGGATCCAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
Db 181 GATTTCCTTTTGGAGGATCTTAAACGATGGCTTTTCAGGACAGGTGTCTCTCCCAAGAA 240
QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACATAATAGTGTACCGATCCC 300
Db 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACAATATAGTGTAAACAGCCT 300
QY 301 TCTGGAAAAATCCTTGTGTTAGATCTCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
```

```
Db 301 TCTGGAATACTCTGCTGTGATCCTCCAGTAATATCCCGGATTAACCTTAATGTAAA 360
Qy 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATATGG 420
Db 361 ACTGTTTCATATATCAAGGTCAAAACCTCATGACAGGGGATTCCTCATTTGTTGG 420
Qy 421 GGAGCATTTTTCTGTATGATCGCATGTCCTCCACAAATGATACCGAGGCAAAATCTTC 480
Db 421 GGGGCATTTTTCTGTATGATCGCTGCTCTACAAATGATACCGAGGCAAGTCTTC 480
Qy 481 ACTGAGGGAATATAGCAGCTATGATGTCATATAGACAGAGTGCACAAATGATTTCTCG 540
Db 481 ACTGAGGGAATATAGCAGCTATGATGTCATATAGACAGAGTTCACAGAAATGATTTTCT 540
Qy 541 CGGCAAGGACAGGATCGCTCATATGATCTGACTTCTACTAATAATATATGGACAAGT 600
Db 541 AGGCAAGGACAGGATGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 601 AGTAACGGAAACGCAACGAAATGACACTGATGATGATGATGATGATGATGATGATGAT 660
Db 601 AGCAATGAAACGACAGAGAAATGATGCGGATGTTTGGCATCTCCAAAGATACAACTCC 660
Qy 661 ACAAGAACCAAAACATGCTCGTCCGTCGCAAAATACCTCCACACTGCGCCACAGCCGTCG 720
Db 661 ACAAAACAAATCAAAACATGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTG 720
Qy 721 GAGATCAAACTCACAGACACCCCAACTGATGCCCAAACTCAATACACGACGCCCAAGC 780
Db 721 AGCTTCACTCTACAAATCTCAATTAATATGCTAATCTGGAACCTATGAAACCAAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGGAGAACGAGAAACCCACACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Qy 841 ACTCTGATGCGGTACCAAGCAAGGCTTTCATCAATGATCCACCTCCCTCCCTCACCA 900
Db 841 ACTCTTAATGTATCTCACTGAACAGAAACAAATCGTCAACAAATATGTCACCTCTTCACTA 900
Qy 901 CAACCAAGCACGCCACAGCAAGGAGAAACACAAACCAATTCCTCCAGATCTGTGACT 960
Db 901 CATCCAGACCTCTCAACATGAGCAAAACAGTACGATCTTCCGACATGCTGTAACT 960
Qy 961 GAACTAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCATTAACACTACCA 1020
Db 961 GAGCAATGGAACCGACCAACACAAACAGCAACGCTCTCAACAAATACTAATACA 1020
Qy 1021 ATCTCTACTAACACACCTCCAAACACACTTTCAGCACTCTCTGACCACTTACAAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGGCAAC 1080
Qy 1081 ACCACCAATGACACACAGACGACAAATCACTGAAATGAGCAACAGTCCGCCCTCG 1140
Db 1081 ATCACCAATATGATACACAGCTGNACTAGACAGAAAGGCAACAAACCAATGCTCAGTTG 1140
Qy 1141 ATAACAACCTGCTCCACGGAATCCACCAACAGCAAGAGCACCCAGCAGCAAAAAA 1200
Db 1141 AACACAACCTAGATCCACAGAAATCCACCAACAGGACAGACACCAACAGCAAC 1200
Qy 1201 GGGCCCGCACAAACGACCAACAGCAAAATAGAGCATTTTCAACAGTCTCTCCGCCAC 1260
Db 1201 AACATCATATGACGACATCAGATATAACAGCAAAACACCCCAAAATTTCTTCCGGAT 1260
Qy 1261 CCAGCTCCAGTGCACAACTCTGTATATTTTTCAGAAAGAGGATCG 1308
Db 1261 TCTAGTCCGACAAACCGCCCTCTCTATATATCTTTAGAAAGAAAGATCG 1308
```

RESULT 13

ABT13456

ID ABT13456 standard; DNA; 2046 BP.

XX

AC ABT13456;

XX

```
DT 30-JAN-2003 (first entry)
XX DNA encoding a chimeric filovirus protein RVN-GP1/GP2.
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX Unidentified.
XX Chimeric.
XX WO200279239-A2.
XX 10-OCT-2002.
XX 31-JAN-2002; 2002WO-US003339.
XX 31-JAN-2001; 2001US-0267522P.
XX (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABU18478.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Disclosure; Page 89-91; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
XX Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;
```

Query Match 42.4%; Score 780; DB 10; Length 2046;

Best Local Similarity 74.8%; Pred. No. 5.2e-223;

Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

```
Qy 1 ATGAGACACATGTTTCTTATCAGTCTTATTAATTCAGGACAAAAATCTCCC 60
Db 1 ATGAGACCAATATTTCTGATGATCTCATTTTATCCAAAGTATAAAACTCTCC 60
Qy 61 ATTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCGGA 120
Db 61 GTTTTAGAAATGCTAGTAAGACCAACCTCAAGATGTAGATTCAAGTGTCTCCG 120
Qy 121 CTCGAGACAGAGAGAGCTCATCTGATGGATTCACACTGAGTGGGCAAAAGTTG 180
Db 121 CTCGAGAGACAGAGAGATGTTTCTGATGGATTTACACTGAGTGGGCAAAAGTT 180
Qy 181 GATTCCCTTTGGAGGATCCAGCGATGGGCTTTTCAGACAGAGTGTCTCCCAAG 240
Db 181 GATTCCCTTTGGAGGATCTAAACGATGGGCTTTTCAGACAGAGTGTCTCCCAAG 240
Qy 241 GTTGTAGTACACAGAGGGGAGAGAGCAAAACATGCTACAATATAGTGTAAAG 300
Db 241 GTTGTAGTATACGAGAGAGAGAGAGCAAAACATGTTACATATAGTGTAAAG 300
Qy 301 TCTGAAAAATCTTCTGTTAGATCTCTCTTACCAACATCCGCTGACTATCCGAA 360
Db 301 TCTGAAAAATCTTCTGCTGCTGATCTCTCCAGTAATATCCGCTGATTTACCT 360
Qy 361 ACTATCCATCATATTCAGAGGTCAAAACCTCATGACAGGGGATCGCCCTTCAT 420
Db 361 ACTGTTTCATATTAAGGTCAAAACCTCATGACAGGGGATTTGCCCTTCATTT 420
```

XX	Marburg virus.
OS	
XX	
XX	Key Location/Qualifiers
FH	97..-2142
FT	/+tag= a
FT	/product= "Marburg virus Ravn strain glycoprotein (GP)"
XX	
XX	US6200959-B1.
PN	13-MAR-2001.
PD	
PX	
PF	04-DEC-1996; 96US-00760615.
PP	
XX	04-DEC-1996; 96US-00760615.
XX	(POWD-) POWDERJECT VACCINES INC.
PA	
XX	
XX	Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
PI	
XX	WPI; 2001-280564/29.
DR	P-PSDB; AAE00708.
DR	
XX	
PT	Inducing an immune response to a Marburg or Ebola virus involves
PT	delivering a genetic vaccine, which contains a genetic construct encoding
PT	antigenic determinants for filovirus, using a particle acceleration
PT	device.
XX	
XX	Claim 6; Col 33-40; 33pp; English.
PS	
XX	The patent discloses a method of inducing an immune response to Marburg
CC	or Ebola virus glycoprotein which involves delivering a genetic vaccine,
CC	containing a genetic construct encoding antigenic determinants for
CC	filovirus, using a particle acceleration device. The genetic vaccine for
CC	filovirus is created by joining a DNA sequence encoding at least a
CC	portion of the filovirus glycoprotein to a promoter effective to promote
CC	transcription of the DNA sequence. This method is useful for inducing
CC	humoral, cell-mediated and secretory immune responses in the treated
CC	individual. The present sequence is a DNA encoding Marburg virus Ravn
CC	strain glycoprotein (GP), which is a type I transmembrane protein. This
CC	sequence is used in the construction of genetic vaccine against
CC	filoviruses
XX	
SQ	Sequence 2247 BP; 734 A; 518 C; 424 G; 571 T; 0 U; 0 Other;
	Query Match 42.0%; Score 773.8; DB 4; Length 2247;
	Best Local Similarity 74.6%; Pred. No. 4e-221;
	Matches 973; Conservative 0; Mismatches 332; Indels 0; Gaps 0
QY	1 ATGAAGACCATCTGTTTCCTTATCAGTCTTATCTTAATCAAGGGACAAAAAATCTCCCC 60
Db	97 ATGAAGACCATATATTTCGTATTAGTCTCATTTTAATCAAAGTATAAAACTCTCCCT 156
QY	61 ATTTTAGAGATAGCTAGTAATAATCAACCCAAAATGTGGATTGGTATGCTCCCGAACT 120
Db	157 GTTTTAGAAATTGCTAGTAACAGGCCAACCTCAAGATGTAGATTTCAGTGTCCTCGGAAC 216
QY	121 CTCGAGAGACAGAAGACGTCCTCATCTGATGGGATTCACACTGAGTGGGC AAAAGTTGCT 180
Db	217 CTCGAAAAGACAGAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGC AAAAGTTGCT 276
QY	181 GATTCCTCCCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db	277 GATTCCTCCCTTTGGAGGCATCTAAACGATGGGCTTCAGGACAGGTGTTCCTCCCAAGAC 336
QY	241 GTTAGATACACAGAGGGGGAGAACCCAAAACATGCTACAATATAGTGTAAACGGATCCC 300
Db	337 GTTAGATATACGGAAGGAGAAGACCCAAAACATGTTTACAAATAAGTGTAAACAGACCCCT 396
QY	301 TCTGGAAATCCTTGCTTTAGATCTCTTACCACCATCGTGACTATCGAAATCGAA 360
Db	397 TCTGGAAATCCTTGCTTGCTGGATCTCTCCAGTAATATCCGGATTAATCGCTTAATGTAA 456

[illegible]

361	Qy	ACTATCCATCATATTTCAAGTCAAACCCCTCATGTCACAGGGATCGCCCTTCATTTATGG	420
457	Db	ACTGTTTCATCATATTTCAAGTCAAACCCCTCATGTCACAGGGATGCGCCTCCATTTGTGG	516
421	Qy	GGAGCAATTTTTCTGTATGATCGCATTTGCCCTCCACAAACATGTACCGAGCAAGTCTTC	480
517	Db	GGGGCATTTTTCTGTATGATCGGTTGGCTCTACAAATGTACGGAGCAAGGTCCTC	576
481	Qy	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGAAGTGCACAAAAATGATTTTCTCG	540
577	Db	ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTCACAGATGATTTTCT	636
541	Qy	CGGCAAGGACAAGGGTACCTCATATAGAACTCTGACTTCTACTAATAAATTTGGACAAGT	600
637	Db	AGGCAAGGACAAGGTATTCGTCACTGAACTTGACCTCCACCAATAAATTTGGACAAGC	696
601	Qy	AGTAA CGGACGCAACGATGACCTGGATGTTTCGGCGCTCTTCAAGATACAATTCT	660
697	Db	AGCAATGAACCGAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAATACAACTCC	756
661	Qy	ACAAAGAACCAACATGTGCTCGGTCGCAAAATACCTCCACCATCTGCCACAGAGCCGCTCG	720
757	Db	ACAAACAATCAACATGCCCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	816
721	Qy	GAGATCAAACTCAAGACA CCCCACCTGATGCGCACCAAACTCAATACCAAGGACCCCAAGC	780
817	Db	AGCATTCACCTCTACAAATACTCAAAATTAATACTGCTAAATCTGGAACATATGAACCCCAAGT	876
781	Qy	AGTGATGATCAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACA	840
877	Db	AGGCACATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	936
841	Qy	ACTTCTGATCGGTCACAAAGCAGGCTTTCATCAACAATGCCCACCTCCCTCACCA	900
937	Db	ACTCTTAATGTAGTCACTGAACAGAACAACTCGTCAACATATTTGTCCACTCTTCACTA	996
901	Qy	CAACCAAGCAGCCACAGCAGGAGGAAACAACAACCAATTCACCAAGATGTGTGACT	960
997	Db	CATCCAAGCACCTCACAAACATGAGCAAAAACAGTACGAAATCCTTCCCGACATGCTGAACT	1056
961	Qy	GAACTAGACAATAATACACAACTGCAACACCGTCCATGCCCTCATACACTACCA	1020
1057	Db	GAGCACAATGGAAACCGACCCCAACAACAACACAGCAACGCTCTCTCAACAATCTAATACA	1116
1021	Qy	ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCACCAATTACAAAAC	1080
1117	Db	ACTCCCACCTTAACACTCTCAAGTACAACTCAGTACTCCTTCCCTCCCAACCCGCAAC	1176
1081	Qy	ACCACCAATGACAAACAACAGAGACCAATCACTGAAATGAGCAAAACGATGCCCCCTCG	1140
1177	Db	ATCACCATAATGATACACAACGCTGCAACTAGCAAGAAAGGACAAAACCAATGCTCAGTTG	1236
1141	Qy	ATAACACCTGCTCCAAACGGGAATCCCAACACAGCAAGAGCACCGACGACGCAAAA	1200
1237	Db	AACACAACCTTAGATCCAACAGAAAAATCCCAACACAGGACAAGACACCAAGACCAACC	1296
1201	Qy	GGCCCCGCCACAACGGCACCAACAACAGCAAAATGAGCATTTTCACAGCTCTCTCCCCCACC	1260
1297	Db	AACATCATCATGACGACATCAGATATATACAGGACAAACACCCCAACAATTTCTTCCGGAT	1356
1261	Qy	CCCAGCTCGACTGCACAACACTCTTTGTATATTTCTAGAAAGACGA	1305
1357	Db	TCTAGTCCGACAAACCCGCCCTCTCTATATCTTTTAGAAGAAACGA	1401

RESULT 15	
ACC71562	
ID	ACC71562 standard; DNA; 6902 BP.
XX	
XX	
AC	ACC71562;
XX	
DT	10-JUL-2003 (first entry)
XX	
XX	

DE	VR6703(pVR1012x/s Marburgdelta TM/h (codon optimised)) plasmid.
XX	
XX	Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW	viral infection; filovirus; circular; cyclic; ds.
XX	
XX	Synthetic.
OS	
XX	WO2003028632-A2.
PN	
XX	10-APR-2003.
PD	
XX	
PF	24-SBP-2002; 2002WO-US030251.
PR	
XX	01-OCT-2001; 2001US-0326476P.
PR	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PA	
XX	Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI	WPI; 2003-371961/35.
XX	
DR	
XX	
XX	New bimodal priming and boosting compositions, useful as viral vaccines,
PT	specifically for eliciting an immune response against a filovirus or a
PT	disease caused by infection with filovirus.
XX	
PS	Claim 1; Page 213-215; 219pp; English.
XX	
CC	The present invention relates to a bimodal priming composition and
CC	boosting composition for priming and boosting an immune response to an
CC	antigen in an individual. The compositions comprise (a) a priming
CC	composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC	encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC	virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC	DNA plasmid selected from AC71521-AC71563 and (b) a boosting
CC	composition comprised of a replication-deficient adenovirus, comprising a
CC	nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC	paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC	epitope-bearing domain, or a replication deficient adenovirus selected
CC	from constructs AC71521-AC71563. The compositions are useful as viral
CC	vaccines, specifically for eliciting an immune response against a
CC	filovirus or a disease caused by infection with filovirus e.g. Ebola
CC	virus infection
XX	
SQ	Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;
Query Match 37.4%; Score 688; DB 10; Length 6902;	
Best Local Similarity 70.5%; Pred. No. 4.5e-195;	
Matches 919; Conservative 0; Mismatches 385; Indels 0; Gaps 0;	
Qy	1 ATCAAGACCAACATGTTTCCTTATCAGTCCTATCTTAATCAAGGGACAAAATAATCTCCCC 60
Db	1923 ATGAGGACCACTGCCTGTTTCATCAGCCTGATCTGATCCAGGGCATCAGACCTGCCC 1982
Qy	61 ATTTTAGATAGTCTAGTATTAATCAACCCCAAAATGTGGATTCTCGGTATCTCCGGAAT 120
Db	1983 ATCCTGGATGCCAGACAAACAGCCCAAGACCTGACACGGTGTGCAGCGGCACC 2042
Qy	121 CTCAGAAACAGAGACGTCATCTGATGGGATTCACTAGTAGTGGGCAAAAAGTTGCT 180
Db	2043 CTGCAGAAACCCAGGACGTCGACCTGATGGGCTTCACTCCCTGACGGCCCAAGAGTGGCC 2102
Qy	181 GATTCCCTTTTGAGGCATCCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
Db	2103 GACAGCCCTCTGGAGGCCACAGAGGTGGCCCTTCAGAACCGGCTGCCCCCAAGAAC 2162
Qy	241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTTCAATATAAGTGTAAACGGATCCC 300
Db	2163 GTGAGTACACCGAGGGCGAGGAGCCAGACCTGCTTACAACATCATCGGTGACCGACCCC 2222
Qy	301 TCTGGAAAAATCCTTGCTGTGTAGATCTCTCTTACCAACATCCGTGACTATCCGAATGCAA 360
Db	2223 AGGGCAAGAGCTGCTGCTGGAGCCCTCCACCAACATCAGGACTACCTCTAAGTGCAG 2282

```
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2283 ACCATTCACCATCCAGGGCCAGAAACCTCATGCGCCAGGGCATCGCCCTGCACCTGTGG 2342
QY 421 GGAGCATTTTCTCTGTATGATCGCATGCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2343 GCGGCTTCTCTGTATGATCGCAGAGATCGCCAGCACCACTGTACAGGGGCAAGGTGTC 2402
QY 481 ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTTCTCG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2403 ACCGAGGGCAACATCGCGCCATGATCGTTAAAGAAGCGTGCACAGATGATCTTCAGC 2462
QY 541 CGGCAAGGACAGGCTACCGTCATGATCTGACTTCTACTAATAAATATTGGACAAGT 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2463 AGGCAGGGCCAGGGCTACAGGCACATGAACCTTGACAGCACCAACAGTACTTGGACCAGC 2522
QY 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTGGCGCTCTTCAAGAAATACAAATTCT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2523 AACACGGCACCCAGACCAACGACACCGGCTGCTTCGGGGCCCTGCAGAGTACAAACAGC 2582
QY 661 ACAAAGAACCACAAATGTGCTCCGTCCTCAAAATPACTCCACCACTGCCACACAGCCCGTCCG 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2583 ACCAAGAACCAGACTGTGGCCCCAGCAAGATCCCCAGCCCTGCCCCACCGCCAGGCC 2642
QY 721 GAGTCAAACTCACAGCACCCCACTGATGSCCACCACCACTCAATACACGGACCCAAAGC 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2643 GAGATCAAGCCACACAGCACCCCCACCGACGCCCCCTGAAACACACCGACCCCAAC 2702
QY 781 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2703 AACGACGAGGAGACTGATCACAGCGGAGCGGAGCGGAGGAGGAGGCTTACACC 2762
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2763 ACCAGCGCGCGTGACCAAGCAGGGGCTTGAGCAGCACCATGCTCTTACCCCTAGCCCT 2822
QY 901 CAACCAAGCACGCCACAGGAGGAGGAACAACAACCAATTCACCAAGATGCTGTGACT 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2823 CAGCCCGACGCCCTCAGCAGGAGGGCAACAACACCGACCAAGCCAGCGGCAACCGTGACC 2882
QY 961 GAACTAGACAAAATAACACAACTGCACAAACCGTCCATGCCCTCATAACTACACACA 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2883 GAGCCCAACAGACCAACACACCCGCCAGCCGATGCTCTTCAACACACACCCGCC 2942
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTGTGCACCAATTAACAAAC 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2943 ATCAGCACCAACAACACAGCAAGAACTTTCAGCACCTTGAGCGTGCAGAAC 3002
QY 1081 ACCACCAATGACACACACAGAGCAATCACTGAAAATGAGCAAAACAGTGCCCCCTCG 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3003 ACCACCAACTAGCAGACACCCAGAGCACCGCCACCGAGAACGAGCAGACGCGCCCTAGC 3062
QY 1141 ATAACAACCTGCTCCCAACGGGAATCCACACAGCAAGAGACCCAGCAGCAABAAA 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3063 AAGACACCTTGCCTCCACCGGCAACTGTGACACCGCAAGAGACCAACACCAACAG 3122
QY 1201 GGCCCCGCCACACAGGACCAACACGACAAATGAGCATTTTCAACAGTCTCTCCCCCACC 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3123 GGCCCCACACACCGCCCTTAACATGACCAACGACCCACCTGACCGCCCCCAGCCCCACC 3182
QY 1261 CCCAGCTGACTGCAACAATCTTGTATATTTTCAAGAAAGCG 1304
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3183 CCCAACCCCAACCCAGCACCTGTGTACTTTCAGGAGAGAG 3226
```

Search completed: July 31, 2005, 16:57:16
Job time : 968.473 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6009.96 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctct.....tttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsel1:*

9: gb_gsel2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	99	5.4	895	9	CNS0071A
C 2	90.4	4.9	1101	9	CNS00FXE
C 3	87.2	4.7	922	9	CNS0073W
C 4	83.4	4.5	884	9	CNS006U0
C 5	80.8	4.4	1225	9	CNS0156K
C 6	79.2	4.3	909	9	CNS000TL
C 7	75.6	4.1	1101	9	CNS000JL
C 8	75	4.0	1965	9	AG136151
C 9	73.6	3.9	939	9	CNS006U0
C 10	72.4	3.9	1101	9	CNS017ZT
C 11	72.2	3.9	1787	9	CG754239
C 12	72.2	3.9	1626	9	AG131986
C 13	71.8	3.9	1455	9	AG382036
C 14	71.4	3.9	1101	9	CNS00LO0
C 15	70.2	3.8	1101	9	CNS017YH
C 16	70	3.8	1101	9	CNS006QP
C 17	69.2	3.8	902	9	CNS00075A
C 18	68.2	3.7	861	9	CNS00075A
C 19	68	3.7	939	9	CNS000CNG
C 20	67.4	3.7	1559	9	CG756460
C 21	67	3.6	796	8	B12592
C 22	67	3.6	993	9	CNS005N6
C 23	67	3.6	1319	9	AG371230
C 24	67	3.6	1667	9	CG756610

C 25	66.8	3.6	913	6	CA986290
C 26	66.2	3.6	470	4	BG786319
C 27	66	3.6	723	8	AZ627893
C 28	66	3.6	1025	9	CL487193
C 29	65.6	3.6	527	9	CE480293
C 30	65.4	3.6	925	7	CK425533
C 31	65	3.5	769	8	BZ579780
C 32	65	3.5	869	9	AG137111
C 33	65	3.5	1147	8	CC187235
C 34	64.8	3.5	513	9	CNS02CMF
C 35	64.8	3.5	633	7	CV070818
C 36	64.8	3.5	860	9	CNS018FL
C 37	64.6	3.5	937	9	CNS006ST
C 38	64.6	3.5	1153	9	CNS07BVP
C 39	64.6	3.5	1345	8	AQ743328
C 40	64.6	3.5	1613	8	BZ557155
C 41	64.4	3.5	917	9	CL466449
C 42	64.2	3.5	1131	9	CNS03AWV
C 43	64	3.5	1350	9	AG393392
C 44	64	3.5	2263	3	CR698468
C 45	63.8	3.5	732	9	AG135778

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS

DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #

ACCSSION
AL066286

VERSION
GSS.

KEYWORDS
GSS.

SOURCE
Drosophila melanogaster (fruit fly)

ORGANISM
Drosophila melanogaster

REFERENCE
1 (bases 1 to 895)

AUTHORS
Genoscope.

TITLE
Direct Submission

JOURNAL
Submitted (02-JUN-1999)

COMMENT
Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammeter in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

1..895

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACRI4B09"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match

Best Local Similarity

5.4%; Score 99; DB 9; Length 895;

24.1%; Pred. No. 6.1e-17;

[illegible]

RESULT 2	CNS00FXE	1101 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	BACR32C19				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL071370				
VERSION	AL071370.1	GI:4951210			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyraidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
source	1..1101 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACR32C19" /clone_1ib="RPCI-98"				

/note="end : TET3"

Query Match 4.9%; Score 90.4; DB 9; Length 1101;
Best Local Similarity 27.5%; Pred. No. 2 le-14;
Matches 108; Conservative 135; Mismatches 149; Indels 1; Gaps 1;

Qy 872 CATCAACAATGCGACCCACTCTCCTCACCAACCAAGCAGCCACAGAGGAGGAACA 931
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db MATAAMAHAYACAAAAAACMMWMAWAAACAAAWACMNMACAMMCMCMWMAAMMM 1037

Qy 932 ACACAAACCATTCCCAGATGTCTGACTGTAAGTACGACAAAATAATACACAACCTGCACAAC 991
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db ACAGMH-MCMRMWMCMAAHANMAWMAWAAWAAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 978

Qy 992 CGTCGATGCCCTCATACACTACCAATCTCTACTACAAACACCTCTCAAACACAACACT 1051
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db MAMMMNINCAAMCWAMAMAMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 918

Qy 1052 TCAGCACTCTCTCTGACCACTTACAAACACCAACCAATGACAAACACACAGAGCACATCA 1111
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db CATACTMTTMMCMWAACMMWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 858

Qy 1112 CTGAAATGAGGAAACGAGTGCCCTCTGATACAAACCTGCTCTCCAACGGGAAATCCCA 1171
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db MMWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 798

Qy 1172 CCACAGAAGACACGACGACGAAAAAGGCCCGCCACAAACGGCACCAACACGACAA 1231
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db MMWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 738

Qy 1232 ATGAGCATTTTCCAGCTCTCCCCCACCACCCCA 1264
||| ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db AMACCCCMCACCCCTCM 705

RESULT 3
CONS0073w/c 922 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL066784
VERSION AL066784.1 GI:4945247
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 922)

REFERENCE Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequences :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..922


```
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : TET3"
```

ORIGIN

```
Query Match      4.7%; Score 87.2; DB 9; Length 922;
Best Local Similarity 22.1%; Pred. No. 1.7e-13;
Matches 78; Conservative 145; Mismatches 130; Indels 0; Gaps 0;

Qy 930 CAACACAAACCATCCCAAGATGCTGCTGACTGACAGCAAAATACACAACTGCACA 989
Db 905 MMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 846
Qy 990 ACCGTCCATGCCCCCTCTAATACACTTACCAATCTCTACTAACACACCTCTCAACACAA 1049
Db 845 MCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 786
Qy 1050 CTTGAGCACTCTCTGTCACCAATTACAAACACCAACCAATGACACACAGAGACAAT 1109
Db 785 MMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 726
Qy 1110 CACTGAAATGAGCAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
Db 725 MMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 666
Qy 1170 CACCACAGCAAGAGCAGCAGCAGCAAAAGAGGCGCCGACCAACGACCAACACGAC 1229
Db 665 AMAMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 606
Qy 1230 AAATGAGCATTTTACCAGTCTCTCCCGCCACCCCGAGCTCGACTGCAACAATC 1282
Db 605 MMAACMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 553
```

RESULT 4

```
CNS006U0/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
```

AL065923

AL065923.1 GI:4944891

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

```
source
1..884
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : T7"
```

ORIGIN

```
Query Match      4.5%; Score 83.4; DB 9; Length 884;
Best Local Similarity 20.1%; Pred. No. 2.1e-12;
Matches 78; Conservative 171; Mismatches 138; Indels 1; Gaps 1;

Qy 855 CACCAAGCAGGCGTTTCATCAACAATGCCACCACTCCCTCACCACAAACCAACGACGCC 914
Db 881 MMYVMNMHTTKKTKTTHMYNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNM 822
Qy 915 ACAGCAGGAGGAGAAACAACAACCAATTCCTCAAGATGCTGTGACTGAATCTAGACAAAA 974
Db 821 MMMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNM 762
Qy 975 TAACACAACTGCACAAACCGTCCATGCGCCCTCATATAACACTACCAACAATCTCTACTAACAA 1034
Db 761 MCMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNMNM 702
Qy 1035 CACCTCAAACACAACTTTCAGCACTCTCTGTGACCACTTACAAACACCAACCAATGACAA 1094
Db 701 MAMM-CMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 643
Qy 1095 CACACAGCAGCACAATCACTGAAATGAGCAAAACAGTCCCTCGATTAACAACCTCTGCC 1154
Db 642 CCCCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 583
Qy 1155 TCACACGGGAAATCCCAACACAGCAGCAAGACAGCAAGCAAAAGAGCGCCCGCCACAAC 1214
Db 582 CAMCTCACCAAMCMCAAHMMAAAAMMAMCCABCMCMCMCMCMCMCMCMCMCMCMCMCM 523
Qy 1215 GGCACCAAAACAGCAAAATGAGCAATTC 1242
Db 522 KCBMCMYBCCCCCCCCCMAMMAMMAAKCKC 495
```

RESULT 5

CNS0166K/c

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence SP6 end of BAC

BACN15C20 of DrosBAC library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

AL063358

AL063358.1 GI:5621512

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBeloBAC11.

FEATURES

Location/Qualifiers

1..1225

/organism="Drosophila melanogaster"

/mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN15C20" /clone_lib="DrosBAC" /plasmid="pBelosBAC11" /note="end : SP6"		Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ogoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	
FEATURES		Location/Qualifiers	
source		1..909	
		/organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACR39G08" /clone_lib="RPCI-98" /note="end : TET3"	
ORIGIN		Query Match 4.3%; Score 79.2; DB 9; Length 909; Best Local Similarity 26.7%; Pred. No. 3.6e-11; Matches 107; Conservative 133; Mismatches 158; Indels 3; Gaps 1;	
QY	832	CCCACACAACTTCTGTATGGGTCCACCAAGCAGGGCTTTCATCAACAATGCCACCCTACT	891
DB	893	CMCCMCCMCC	834
QY	892	CCCTCACACACCAAGCAGCCACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	951
DB	893	CMCRCMCCM	774
QY	952	GCTGTGACTGAATCTAGACAAATAAACAACAATGCGCAACCGTCCATCGCCCTCATAC	1011
DB	773	AMAMAMMMMMMMMAAMMAAACAAMMAACMACCCCMCMCMCMMAAMMAAMMAAMMAAM	714
QY	1012	ACTACCACAATCTCTACTAAACAACCTCCAAACAACAATTCAGACTCTCTCTCTGACCA	1071
DB	713	AMMAAAMAAACM	654
QY	1072	TTACAAAACACCAATATGACACACACGACGACGACATCACTGAAATGAGCAAAACGCT	1131
DB	653	CMCCACAAAMMAAMMAAACAAMCAACMACCMCMCMCMCMCMCMCMCMCMCMCMCMCM	594
QY	1132	GCCTCCTCGATAACAACCTCTCCAAACGGGAAATCCCAACGACGACGACGACGACGACG	1191
DB	593	MAAMCMCCCM	537
QY	1192	AGCAAAAAGGCGCCGACCAACGCGCACCAACCAACGACGACGACGACGACGACGACG	1232
DB	536	ACMAAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAM	496
RESULT 7		CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999	
LOCUS		Drosophila melanogaster genome survey sequence TET3 end of BAC;	
DEFINITION		BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
ACCESSION		AL078714	
VERSION		AL078714.1 GI:5102004	
KEYWORDS		GSS.	
SOURCE		Drosophila melanogaster (fruit fly)	
ORGANISM		Drosophila melanogaster	
AUTHORS		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
TITLE		1 (bases 1 to 1101)	
JOURNAL		REFERENCE 1	
BP 191 91006		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
WWW.GENOSCOPE.CNS.FR		REFERENCE 1 (bases 1 to 1101)	

```

AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazutoyo Osoegawa and
              Aarton Mammoser in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCL-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain y2; cn bw sp, the same strain used for the BDGP's
              P1 and EST libraries. A more detailed description of the library
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES     source
              1..1101
                /organism="Drosophila melanogaster"
                /mol_type="genomic DNA"
                /db_xref="taxon:7227"
                /clone="BACR48P19"
                /clone_lib="RPCL-98"
                /note="end : TET3"

ORIGIN
Query Match      4.1%; Score 75.6; DB 9; Length 1101;
Best Local Similarity 13.8%; Pred. No. 4.3e-10;
Matches 60; Conservative 208; Mismatches 166; Indels 0; Gaps 0;

QY 868 CTTTCATCAACCAATGCCACCCACTCCCTCACCAACCAAGCAGCCACAGCAAGGGA 927
Db 1096 HHTTMMMMCHHTAAMMMAMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 1037

QY 928 AACACACAAACCAATTCACAGATCTGTGACTGAATAGACAAATAACACAACTGCA 987
Db 1036 THMMMMMMMMMMAMMMAMMMMMMMMAATTTTTHMMAMMAVHTTMMMMTMMMAAATTTHAHTT 977

QY 988 CAACCGTCATGCCCTCATTAACACTACCACATCTCTACTAACAACACCTCCAAACAC 1047
Db 976 TTHTTTMMAMCMTHTTMMMMMMMMMMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCM 917

QY 1048 AACTTCAGCACTCTCTGACACATTACAAACACACCAATGACACACACAGAGACA 1107
Db 916 MMTTTHHHMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMMAA 857

QY 1108 ATCACTGAAATGAGCAACAGTCGCCCTCGATAACAACCTCGCTCCAAACGGGAAT 1167
Db 856 MMAMMMMMCMHMMAMMTTTTTHMMAMMAVHTTMMMMTMMMMMMCMHMMMAAATTMM 797

QY 1168 CCCACACAGCAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1227
Db 796 MMAMAMHTTHHCTTMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 737

QY 1228 ACAATGACATTTTACCACTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1287
Db 736 TTMMHMMMAATMTTMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMM 677

QY 1288 TATTTTCAGAGAAA 1301
Db 676 TAATAAAAAA 663

RESULT 8
AG136151
LOCUS      AG136151
DEFINITION Pan troglodytes DNA, clone: PTB-149K08.F, genomic survey sequence.
ACCESSION AG136151
VERSION   AG136151.1 GI:16665829

```

```

KEYWORDS     GSS.
SOURCE       Pan troglodytes (chimpanzee)
ORGANISM     Pan troglodytes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE    1
AUTHORS      Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
              Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE        BAC end sequences of library PTB
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 837)
AUTHORS      Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
              Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE        Direct Submission
JOURNAL      Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
              1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              (E-mail: chimbes@gsc.riken.go.jp, URL: http://bgp.gsc.riken.go.jp/,
              Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT      Clones are derived from the chimpanzee BAC library PTB. This BAC end
              was generated during the R&D process and may have higher chance of
              clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY      Vector : pKS145
              R.Site 1 : SacI
              R.Site 2 : SacI.
FEATURES     source
              1..837
                /organism="Pan troglodytes"
                /mol_type="genomic DNA"
                /db_xref="taxon:9598"
                /clone="PTB-149K08.F"
                /sex="male"
                /cell_type="lymphoblast"
                /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      4.1%; Score 75; DB 9; Length 837;
Best Local Similarity 47.1%; Pred. No. 5.9e-10;
Matches 264; Conservative 0; Mismatches 292; Indels 5; Gaps 1;

QY 727 AAATCTCAAGCACCCCACTGATGCGACCAAACTCAATACCGAGCCCAAGCGATGAT 786
Db 261 AAACCAACCCACCCACCAACCAACCAACCAACCACTTATCCACCCACCCACCCACA 320

QY 787 GATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAACGAGAACCCACCACTTCT 846
Db 321 AACACACACACCAAAACACACACCAACCAACCAACCAACCAACCAACCAACCAAC 380

QY 847 GATGGGTGACCAAGCAAGGCTTTTCATCAACAAATGCGACCCACTCTCCCTCACCACCA 906
Db 381 AACACCAACCAACCAAAACCTTTCTCAGCGACCAACCAACCAACCAACCAACCAAC 440

QY 907 AGCAACGACAGCAAGGAGGAAAACAACAACCAATTTCCCAAGATGTGTGACTGAAC 966
Db 441 AGCAACCAACCAACCAAAACACACACAAACAAACAAACAAACCAACCAACCAACCA 500

QY 967 GA-----CAAAATAACACAACTGCAACACGTCATGCGCCCTCATTAACACTACCA 1021
Db 501 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 560

QY 1022 TCTGTACTTAACAACCTCCCAACCAACCACTTTCAGCACTCTCTCTGACCACTTACAAA 1081
Db 561 CCACAAACCAACCAATCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 620

QY 1082 CCACCAATGACACACACAGAGCAAACTCACTGAAAATGAGCAAAACAGTGCCTCGA 1141
Db 621 CCAACACAATCAACACAAACCAAAACAAATCAACAAACCAAAACCAACCAACCAACCA 680

QY 1142 TAACACCTCTCTTCCACGGGAATCCCAACCAACGAGCAAGGACGACGACGACGAA 1201

```

```
Db 681 AACAACCTCAACCAACCAACCAAAATCCAAAACACAAACAAACCAACACCA 740
Qy 1202 GCCCGCCACAAACGGCACCACAAACAGACAAATGAGCATTTTACACGAGTCTCCCGCCACCC 1261
Db 741 ACCAACCCCGCCCAACACACACCAACCAACCAACCAACCAACCAACCAACCA 800
Qy 1262 CCAGTCGACTGCACACATC 1282
Db 801 AAACCAAAACAAACCAACCAACC 821

RESULT 9
LOCUS CL090566/c
DEFINITION ISB1-17017 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-17017,
genomic survey sequence.
ACCESSION CL090566
VERSION CL090566.1 GI:40584201
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1965)
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 781
High quality sequence stop: 832.

FEATURES
source
1..1965
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-17017"
/clone_lib="ISB1"
/notes="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN
Query Match 4.0%; Score 73.6; DB 9; Length 1965;
Best Local Similarity 48.8%; Pred. No. 2e-09;
Matches 199; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

Qy 875 CAACAAATGCCACCCACTCCCTCACCACAAACCAAGCAGCCACGAGGAGGAACACA 934
Db 1185 CACCAACACAAACACCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCC 1126
Qy 935 CAAACATTTCCCAAGATGCTGTGACTAGACAGAAATAACAACACTGCACACCGT 994
Db 1125 ACACCACCAACCAACCCACCCCAACCAACCAACCAACCAACCAACCAACCAAC 1066
Qy 995 CCATGCCCCCTCATACACTAGCAATCTCTACTAAACAAACACCTCCAAACACAACTTCA 1054
Db 1065 CCCCCACACACCCACACCCACACCAACAAACCAAGCCACCCACCCACCAACCAACCA 1006
Qy 1055 GCATCTCTCTGGACCATTAACAAAACACCAACCAATGACAAACACAGAGCACAATCTG 1114
Db 1005 CCCCACCAACAAACCCACCCACCCCAACCAACCAACCAACCAACCAACCAACCA 946
Qy 1115 AAAATGAGCAACCAACGAGTGCCTCGATAAACACCTGCCTCAACGGGAATCCACCA 1174
Db 945 CACACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 886
```

```
Qy 1175 CAGCAAGAGGAGCCAGCAGCAAAAAGGCCCCGCCACAAACGGCACCACCAACAAATG 1234
Db 885 CAACACGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 826
Qy 1235 AGCATTTTACCAGTCTCTCCCGCCACCCACCCAGCTGCAGTGCACACATC 1282
Db 825 AAAAAAAACACACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 778

RESULT 10
LOCUS CNS006UU/c
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14L09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL065953
VERSION AL065953.1 GI:4944921
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 939)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oseegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw ap, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..939
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14L09"
/clone_lib="RPCI-98"
/notes="end : T7"

ORIGIN
Query Match 3.9%; Score 72.6; DB 9; Length 939;
Best Local Similarity 33.8%; Pred. No. 3.1e-09;
Matches 124; Conservative 92; Mismatches 150; Indels 1; Gaps 1;

Qy 795 CCTCGCAACATCCGCTCAGGTCGGAGACGAGNACCCCAACCACTTCTGTGCGGT 854
Db 865 MSSCMSACCAACRCMGMAAGCCCGCCCRAGACGAGACRGMGASVCGVASVGVCA 806
Qy 855 CACCAAGCAAGGGTTTTCATCAAAATGCCACCACTCCCTCACCAACCAAGCAGCC 914
Db 805 VGSVRMACGAGSAGVSGVRCACACGVMVVCASCSMACAMMARMCAMAVASAAANAA 746
Qy 915 ACAGCAGGAGGAACAAACACAAACCAATTCCTCCAGATGCTGTGACTGACTAGACAAA 974
Db 745 VAHVAGAVARGGGRAAGCGCAACRMAACMWCACAMAMARMSACGSAACAMMAACAS 686
Qy 975 TAAACAACACTGCACACCGTCCATGCGCCCTCATACACTACCACTACCAACTCTACTAACAA 1034
Db 685 VARMACAMAMACAMMAAAACSMRCVMVCVSRACASMACCAACCAACACRCCCMCMCA 626
```

QY	1112	CTGAAATGAGCAACCGAGTCGCCCTCGATATACAACCCCTGCCTCACAACGGGAATCCCA	1171
Db	860	CCACNCACMACAACAAAACAAAAAAMAAAAAAMAWCMCAHMAACCCCCCACMAAAACMMMC	801
QY	1172	CCACAGCAAAGAGCACGACGAGCAAAAAGAGGCCGCCACACGCGCACCAACACGACAA	1231
Db	800	CCCMCAMACMCMYAMCAWCHYCITTTTTTTTNTACHYYCMAAHMAYAYATTACY	741
QY	1232	ATGAGCATTTCCAGAGTCCTCCCCCACCAGCTCGACTGCACAACATCTTGTAAT	1291
Db	740	CAAAWAVYTYYYMYCYCTCTHMVATACVMCCYAHYVWTYYYWTMTWTHAWAMTW	681
QY	1292	TCAAGAAGAGCGATCGGCAATTGTCAATGCTCAACCCAAATCGCAACCCCTTAATTACAT	1351
Db	680	YTATWCMCAAMCTMMCAAWTCATCMCCCTCCTTCYCYYYYYTACMYACCTKCCYT	621
QY	1352	ACT 1354	
Db	620	YCT 618	
RESULT 12			
CG754239/c			
LOCUS			
DEFINITION			
PG49-3-DL2.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,			
genomic survey sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Pristionchus pacificus			
Pristionchus pacificus			
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;			
Neodiplogasteridae; Pristionchus.			
1 (bases 1 to 1787)			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PubMed			
COMMENT			
Contact: Sommer RJ			
Evolutionary Biology			
Max-Planck-Institute for Developmental Biology			
Spemannstr. 37-39, Tuebingen D-72076, Germany			
Tel: 00497071601371			
Fax: 00497071601498			
Email: ralf.sommer@tuebingen.mpg.de			
Class: BAC ends.			
Location/Qualifiers			
1..1787			
/organism="Pristionchus pacificus"			
/mol_type="genomic DNA"			
/strain="California"			
/db_xref="taxon:54126"			
/clone_lib="Ppa EcoRI BAC Library"			
/note="The library was generated by a partial digest of			
the genomic DNA with EcoRI and cloning into the BAC			
vector."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 194; Conservative			
Score 72.2; DB 9; Length 1787;			
Pred. No. 5e-09;			
Mismatches 203; Indels 0; Gaps 0;			
QY	883	CCACCCACTCCCTCACCAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAT	942
Db	758	CCACCAACCATATCACACCACACACACCCCTCCCATCAACACAAACACACACACCA	699
QY	943	TCCAAGATGCTGTGACTGAACCTAGACAAAAATAACAACTGCACACCGTCATGCC	1002
Db	698	ACACCAACAGAGACGACCTTAACACACAAACACCCCACCAAAAAAAAAAGCCACCCC	639

```

QY 1003 CCTCATACACTTACCAATCTCTACTAACACACCTCCAAACAACTTTCAGCACTCTC 1062
Db 638 AACACCCCCACACACACCAAAACAAACAAACCAACCACTACACAAACACAC 579
QY 1063 TCTGCACCATTAACAAAACACCAACCAATGACACACACAGAGCACAATCACTGAAATGAG 1122
Db 578 TCAACACATCAACACATACACACACCCCTACACTTCACACCCACACCCAC 519
QY 1123 CAACACAGTGGCCCCCTCGATTAACAACCTTGCTCCAAACGGGAAATCCCAACCAAGAAAG 1182
Db 518 AACACACAAACAAACACCAACCCCAACCCCAACCAACCCCAACCAACATCCCAAC 459
QY 1183 AGCACACAGCAAAAAGGCCCCCAGCAACAGCGGACCAACACACACAAATGAGATTTC 1242
Db 458 ACCACCATCCCAACACACACACACACCCCAACCAACCAACCAACCAACCAACCA 399
QY 1243 ACCAGTCTCCCCCAACCCCACTGCTGAGTGCACAAAC 1279
Db 398 ACNACCCCAACCAACACACACACACACACACACACACAC 362

RESULT 13
AG131986
LOCUS Pan troglodytes DNA, clone: PTB-144D13.R, genomic survey sequence.
ACCESSION AG131986
VERSION AG131986.1 GI:16661664
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 1626)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : PKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .1626
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-144D13.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 3.9%; Score 71.8; DB 9; Length 1626;
Best Local Similarity 39.7%; Pred. NO. 6.3e-09;
Matches 238; Conservative 0; Mismatches 362; Indels 0; Gaps 0;

QY 647 AGAATACAAATTCACAAAGACCAACATGTGCTCGTCCAAATATACCTCCCACTGC 706
Db 156 AAANNAACAAACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 215

```

```

QY 707 CCACAGCCCGTCCGGAGATCAAACTCACAAGCACCCCAACTGATGTCACCAAACTCAATA 766
Db 216 CCACANANAAACNNNAACCAANNAACCCNCCACCCNCCNNAACCAACCAACCAACCA 275
QY 767 CCACGGACCCCAAGCAGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAAC 826
Db 276 NANCNCNCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 335
QY 827 GAGAACCCCAACACAACTTCTTGATGGGTACCAAGCAAGGGCTTTTCATCAACAATGCCAC 886
Db 336 NNNNNCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 395
QY 887 CCATCTCCTTACCAACCAACGACGCGCACAGCAGGAGGAGAAACACACAAACCACTCCC 946
Db 396 CCCAAACCCCAACACCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 455
QY 947 AAGATGTGTGACTGAATAGACAAACAAATTAACACAACTGCAACACCCGTCATGCCCTCTC 1006
Db 456 CACANACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 515
QY 1007 ATAACACTTACCAATCTTACTAACACACCTCCAAACACAACTTCAGCACTTCTCTG 1066
Db 516 NACACACACCCACANANAAACCAAAACACAAACCAACANANAAACCAACCAACCAACCA 575
QY 1067 CACATTACAAACACCAACCAATGACACACACAGACACCAATCACTGAAATGAGCAAA 1126
Db 576 CACAAACCCCAACACCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 635
QY 1127 CCAGTGGCCCTCGATTAACAACTGCTCCAAAGGGAAATCCCAACACAGCAAGAGCA 1186
Db 636 CCCACCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 695
QY 1187 CCAGCAGCAAAAAGGCGCCGCAACACCGGACCAACACCAACCAACCAACCAACCAAC 1246
Db 696 ACCNCCNCAACCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 755

RESULT 14
AG382036/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-192P03.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG382036
VERSION AG382036.1 GI:47993241
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1455)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunuya Abe (abe@tc.riken.jp).
Taikuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@tc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBac3.6
Vector : pBac3.6

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 294.875 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctc.....tttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303.4	70.8	2164	3	US-08-760-615-3
2	1303.4	70.8	11460	4	US-09-336-910A-1
3	773.8	42.0	2247	3	US-08-760-615-5
4	526	28.6	2298	4	US-09-650-086A-1
5	525.6	28.5	2172	3	US-08-760-615-1
6	74.6	4.1	152132	4	US-09-949-016-13845
7	74.6	4.1	152145	4	US-09-949-016-12371
8	60.4	3.3	34230	4	US-09-949-016-12052
9	60.4	3.3	128470	4	US-09-949-016-13765
10	57.6	3.1	5511	3	US-08-928-361B-2
11	57.6	3.1	5511	4	US-09-588-995A-2
12	57.6	3.1	7334	3	US-08-928-361B-1
13	57.6	3.1	7334	4	US-09-588-995A-1
14	54.2	2.9	30656	4	US-09-949-016-14613
15	54	2.9	522	4	US-09-949-016-103758
16	53.6	2.9	5163	3	US-08-700-651-1
17	53.6	2.9	5163	3	US-08-928-361B-4
18	53.6	2.9	5163	4	US-09-588-995A-4
19	53.6	2.9	5318	3	US-08-700-651-2
20	53.6	2.9	5318	3	US-08-928-361B-3
21	53.6	2.9	5318	4	US-09-588-995A-3
22	52.6	2.9	271134	4	US-09-949-016-12705
23	52.6	2.9	305491	4	US-09-949-016-17550
24	52.4	2.8	39154	4	US-09-949-016-12384
25	52.4	2.8	39154	4	US-09-949-016-12801
26	52.4	2.8	39443	4	US-09-949-016-14326
27	52.4	2.8	39443	4	US-09-949-016-14327

28	51.6	2.8	114793	4	US-10-148-806-3	Sequence 3, Appli
29	51	2.8	57280	4	US-09-949-016-11796	Sequence 11796, A
30	51	2.8	57280	4	US-09-949-016-12843	Sequence 12843, A
31	51	2.8	57280	4	US-09-949-016-12844	Sequence 12844, A
32	51	2.8	57280	4	US-09-949-016-12846	Sequence 12846, A
33	51	2.8	57280	4	US-09-949-016-13542	Sequence 13542, A
34	51	2.8	57280	4	US-09-949-016-13543	Sequence 13543, A
35	51	2.8	57280	4	US-09-949-016-13544	Sequence 13544, A
36	51	2.8	57280	4	US-09-949-016-13545	Sequence 13545, A
37	51	2.8	57280	4	US-09-949-016-14633	Sequence 14633, A
38	51	2.8	57280	4	US-09-949-016-14634	Sequence 14634, A
39	51	2.8	57280	4	US-09-949-016-14635	Sequence 14635, A
40	51	2.8	57280	4	US-09-949-016-14636	Sequence 14636, A
41	51	2.8	57280	4	US-09-949-016-14637	Sequence 14637, A
42	51	2.8	57280	4	US-09-949-016-14638	Sequence 14638, A
43	51	2.8	57280	4	US-09-949-016-14639	Sequence 14639, A
44	51	2.8	57280	4	US-09-949-016-14640	Sequence 14640, A
45	48.8	2.7	60990	4	US-09-949-016-14080	Sequence 14080, A

ALIGNMENTS

RESULT 1
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3

Query Match		70.8%;	Score 1303.4;	DB 3;	Length 2164;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 1304;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTTCAAGGGACAAAAATCTCCCG	60						
DB	119	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTTCAAGGGACAAAAATCTCCCG	178						
QY	61	ATTTTAGAGTAGTAAATCAACCCCAAAATGTTGGATCGGTATGCTCCGGAAT	120						
DB	179	ATTTTAGAGTAGTAAATCAACCCCAAAATGTTGGATCGGTATGCTCCGGAAT	238						
QY	121	CTCCAGAACAGAACGCTCCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	180						
DB	239	CTCCAGAACAGAACGCTCCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	298						
QY	181	GATTCCTTTTGGAGGATCCAGGATGGGCTTTTCAGGACAGGTGACTCCCAAGAAAT	240						
DB	299	GATTCCTTTTGGAGGATCCAGGATGGGCTTTTCAGGACAGGTGACTCCCAAGAAAT	358						
QY	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGATCCC	300						
DB	359	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGATCCC	418						
QY	301	TCTGGAATAATCTTGTGTAGATCCTCTACCAATCCGTGACTATCCGAAATGCAAA	360						
DB	419	TCTGGAATAATCTTGTGTAGATCCTCTACCAATCCGTGACTATCCGAAATGCAAA	478						
QY	361	ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG	420						
DB	479	ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG	538						
QY	421	GGAGCATTTTCTGTATGATGCGATGCTCCACAAATGATGACCGAGGCAAGTCTTC	480						
DB	539	GGAGCATTTTCTGTATGATGCGATGCTCCACAAATGATGACCGAGGCAAGTCTTC	598						
QY	481	ACTGAAGGAAACATAGCAGCTATGTTGTCATTAAGACAGTGCACAAATGATTTCTCG	540						
DB	599	ACTGAAGGAAACATAGCAGCTATGTTGTCATTAAGACAGTGCACAAATGATTTCTCG	658						
QY	541	CGGCAAGGACAAAGGTACCGTCAATATGAACTGACTTCTTACTAATAAATATTTGCAAGT	600						
DB	659	CGGCAAGGACAAAGGTACCGTCAATATGAACTGACTTCTTACTAATAAATATTTGCAAGT	718						
QY	601	AGTAACGGAAACGAAACGATGACATGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	660						
DB	719	AGTAACGGAAACGAAACGATGACATGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	778						
QY	661	ACAAAGAACCAACATGTGCTCGTCCAAATATACCTCCACCACTGCCACAGCCGCTCG	720						
DB	779	ACAAAGAACCAACATGTGCTCGTCCAAATATACCTCCACCACTGCCACAGCCGCTCG	838						
QY	721	GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCAAGC	780						
DB	839	GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCAAGC	898						
QY	781	AGTGATGATGAGGACCTCCCAACATCCGCTCAGGCTCGGAGAACGAGAACCCACACA	840						
DB	899	AGTGATGATGAGGACCTCCCAACATCCGCTCAGGCTCGGAGAACGAGAACCCACACA	958						
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA	900						
DB	959	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA	1018						
QY	901	CAACCAAGCACCGCACAGAGGAGGAAACAAACAAACCATTTCCCAAGATCTGTGACT	960						
DB	1019	CAACCAAGCACCGCACAGAGGAGGAAACAAACAAACCATTTCCCAAGATCTGTGACT	1078						
QY	961	GAACTAGACAAAATTAACCAACTGACAAACCGTCCATGCCCTCATTAACACTACCACA	1020						
DB	1079	GAACTAGACAAAATTAACCAACTGACAAACCGTCCATGCCCTCATTAACACTACCACA	1138						
RESULT 2									
US-09-336-910A-1									
; Sequence 1, Application US/09336910A									
; Patent No. 6517842									
; GENERAL INFORMATION:									
; APPLICANT: United States Army Medical Research Institute of									
; APPLICANT: Infectious Diseases									
; APPLICANT: Hevey, Michael C.									
; APPLICANT: Negley, Diane L.									
; APPLICANT: Pushko, Peter									
; APPLICANT: Smith, Jonathan F.									
; APPLICANT: Schmaljohn, Alan L.									
; TITLE OF INVENTION: Marburg Virus Vaccines									
; FILE REFERENCE: Army 143									
; CURRENT APPLICATION NUMBER: US/09/336,910A									
; PRIOR FILING DATE: 1999-06-21									
; PRIOR APPLICATION NUMBER: US 60/091,403									
; NUMBER OF SEQ ID NOS: 7									
; SOFTWARE: Word, Microsoft Office 97, IBM compatible									
; SEQ ID NO 1									
; LENGTH: 11460									
; TYPE: DNA									
; ORGANISM: Marburg Virus									
US-09-336-910A-1									
Query Match		70.8%;	Score 1303.4;	DB 4;	Length 11460;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 1304;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTTCAAGGGACAAAAATCTCCCG	60						
DB	5940	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTTCAAGGGACAAAAATCTCCCG	5999						
QY	61	ATTTTAGAGTAGTAAATCAACCCCAAAATGTTGGATCGGTATGCTCCGGAAT	120						
DB	6000	ATTTTAGAGTAGTAAATCAACCCCAAAATGTTGGATCGGTATGCTCCGGAAT	6059						
QY	121	CTCCAGAACAGAACGCTCCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	180						
DB	6060	CTCCAGAACAGAACGCTCCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	6119						
QY	181	GATTCCTTTTGGAGGATCCAGGATGGGCTTTTCAGGACAGGTGACTCCCAAGAAAT	240						
DB	6120	GATTCCTTTTGGAGGATCCAGGATGGGCTTTTCAGGACAGGTGACTCCCAAGAAAT	6179						
QY	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGATCCC	300						
DB	6180	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGATCCC	6239						
QY	301	TCTGGAATAATCTTGTGTAGATCCTCTACCAATCCGTGACTATCCGAAATGCAAA	360						

Db 6240 TCTGGAAATCCTTGCTGTAGATCCTCTACCAACATCCGTGACTATCCTAAATGCAA 6299
Qy 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACACAGGGGATGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGACACAGGGGATGCCCTTCATTTATGG 6359
Qy 421 GGAGCATTTTTCTGTATGTCGATCGCTCCCAACAAGTACCGAGGCAAAATCTTC 480
Db 6360 GGAGCATTTTTCTGTATGTCGATCGCTCCCAACAAGTACCGAGGCAAAATCTTC 6419
Qy 481 ACTGAAGGAAACATAGCAGCTATGATGTCATATAGACAGTGCACAAAATGATTTCTCG 540
Db 6420 ACTGAAGGAAACATAGCAGCTATGATGTCATATAGACAGTGCACAAAATGATTTCTCG 6479
Qy 541 CGGCAAGGAAAGGTACCGTCATATGATTCGATTCCTCTATATATATATTTGCAAGT 600
Db 6480 CGGCAAGGAAAGGTACCGTCATATGATTCGATTCCTCTATATATATATTTGCAAGT 6539
Qy 601 AGTACGGAAACGCAACCAATGACACTGGATGTTTCGGGGCTCTTCAAGAATACAAATCT 660
Db 6540 AGTACGGAAACGCAACCAATGACACTGGATGTTTCGGGGCTCTTCAAGAATACAAATCT 6599
Qy 661 ACAAGAAACCAAAATGCTCCGTCCTCAAAATACCTCCACCACTGCCCAACAGCCGCTCG 720
Db 6600 ACAAGAAACCAAAATGCTCCGTCCTCAAAATACCTCCACCACTGCCCAACAGCCGCTCG 6659
Qy 721 GAGTCAAACTCAAGCAACCCCAACTGATGTCACCAAACTCAATPACCAAGCAACCAAGC 780
Db 6660 GAGTCAAACTCAAGCAACCCCAACTGATGTCACCAAACTCAATPACCAAGCAACCAAGC 6719
Qy 781 AGTGATGATGAGGACTCGACATCCGCTCCGCTCAGGTCGGGAGACGAGACCCCAACA 840
Db 6720 AGTGATGATGAGGACTCGACATCCGCTCCGCTCAGGTCGGGAGACGAGACCCCAACA 6779
Qy 841 ACTTCTGATGCGGTCAACAAAGAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACA 900
Db 6780 ACTTCTGATGCGGTCAACAAAGAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACA 6839
Qy 901 CAACCAAGCAGCCCAAGAGGAAACCAACAAACCAATTCCTCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCCAAGAGGAAACCAACAAACCAATTCCTCAAGATGCTGTGACT 6899
Qy 961 GNACTAGCAAAATTAACACACTGCAACCGTTCATCGCTCCCTCCCTCATAACACTACCA 1020
Db 6900 GNACTAGCAAAATTAACACACTGCAACCGTTCATCGCTCCCTCCCTCATAACACTACCA 6959
Qy 1021 ATCTCTACTTAAACACACTCCAAACCAACCTTCAGCACTCTCTGACCACTTACAAAC 1080
Db 6960 ATCTCTACTTAAACACACTCCAAACCAACCTTCAGCACTCTCTGACCACTTACAAAC 7019
Qy 1081 ACCACCAATGACACACAGAGCAATCACTGAAATGAGCAAAACAGTGCCCTCTCG 1140
Db 7020 ACCACCAATGACACACAGAGCAATCACTGAAATGAGCAAAACAGTGCCCTCTCG 7079
Qy 1141 ATAAACACCTGCTCCCAAGGAAATCCCAACACAGCAAGAGACCAAGCAAGCAAA 1200
Db 7080 ATAAACACCTGCTCCCAAGGAAATCCCAACACAGCAAGAGACCAAGCAAGCAAA 7139
Qy 1201 GGCCTCCGCAACAGGCAACCAACAGCAATGAGCATTTCCAGCTCTCCCTCCAC 1260
Db 7140 GGCCTCCGCAACAGGCAACCAACAGCAATGAGCATTTCCAGCTCTCCCTCCAC 7199
Qy 1261 CCCAGCTCGACTGCAACATCTTGTATATTTTCAAGAAAGCGA 1305
Db 7200 CCCAGCTCGACTGCAACATCTTGTATATTTTCAAGAAAGCGA 7244

RESULT 3

US-08-760-615-5

; Sequence 5, Application US/08760615

; Patent No. 6200959

; GENERAL INFORMATION:

; APPLICANT: Haynes, Joel R

; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Ravn
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..2142
; US-08-760-615-5

Query Match 42.0%; Score 773.8; DB 3; Length 2247;

Best Local Similarity 74.6%; Pred. No. 5e-243;

Matches 973; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 60

Db 97 ATGAAGACCAATATTTTCTGATTAGTCTATTTTAAATCCAAAGTATATAAACTCTCCCT 156

Qy 61 ATTTTAGAGTAGCTAGTAAATCAACCCCAATGCGATTGCGTATGCTCCGGAAT 120

Db 157 GTTTTAGAAAATTTAGTAAACAGCAACCTCAAGATGTAGTGTCTCCGGAAC 216

Qy 121 CTCAGAAGACAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGCAAAAAGTTGCT 180

Db 217 CTCGAAGACAGAGAGTGTTCATCTGATGGGATTTACACTGAGTGGCAAAAAGTTGCT 276

Qy 181 GATTCCCTTTGGAGGATCCAGCGATGCGCTTTCAAGACAGGTGTACTCCCAAGAT 240

Db 277 GATTCCCTTTGGAAGCATCTAAACGATGGGCTTTCAAGACAGGTGTCTCCCAAGAAC 336

Qy 241 GTTGAGTACAGAGGGGAGAGCCCAACATGCTACATATAGTAAAGTAAACGGTCCC 300

Db 337 GTTGAGTATACGGAAGGAGAGAGCCCAACATGTTTACATATAGTAAAGTAAACGGTCCC 396

Qy 301 TCTGGAATATCTTGTGTTAGATCTCTCTACCAACATCCGCTACTATCCGAAATGCAAA 360

Db 397 TCTGGAATATCTTGTGTTAGATCTCTCTACCAAGTAAATATCCCGGATTAACCTTAATGTANA 456

ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229.91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ebola virus
STRAIN: Zaire
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2172
OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1

Query Match 28.5%; Score 525.6; DB 3; Length 2172;
Best Local Similarity 98.3%; Pred. No. 1.9e-161;
Matches 531; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1294 AGAAGAAACGATCGGAAATGTGCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1353
Db |||||
Qy 1633 AGAATCGAAGAGAGCAATGTGCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1692
Db |||||
Qy 1354 TGGACTACTCAGGATGAAGGTGCTCAATCGGCTCGACTGATACCATATTTTCGGGCCA 1413
Db |||||
Qy 1693 TGGACTACTCAGGATGAAGGTGCTCAATCGGCTCGACTGATACCATATTTTCGGGCCA 1752
Db |||||
Qy 1414 GCAGCCGAGGGAATTTACATAGAGGGGCTAATGCAATCAAGATGGTTTAACTCTGTGGG 1473
Db |||||
Qy 1753 GCAGCCGAGGGAATTTACATAGAGGGGCTAATGCAATCAAGATGGTTTAACTCTGTGGG 1812
Db |||||
Qy 1474 TTGAGACAGCTGGCCAAAGACGACGACTCAAGCTTTCAACTGTTCTCGAGGCCACAAT 1533
Db |||||
Qy 1813 TTGAGACAGCTGGCCAAAGACGACGACTCAAGCTTTCAACTGTTCTCGAGGCCACAAT 1872
Db |||||
Qy 1534 GAGCTACCCACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTCGCGGATGG 1593
Db |||||
Qy 1873 GAGCTACCGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTCGCGGATGG 1932
Db |||||
Qy 1594 GCGGCACATGCGCAATTTCTGGGACCGGACTGCTGTATCGAACCACATGATTTGACCAAG 1653
Db |||||
Qy 1933 GCGGCACATGCGCAATTTCTGGGACCGGACTGCTGTATCGAACCACATGATTTGACCAAG 1992
Db |||||
Qy 1654 AACATAACAGACAAATTTGATCAGATTAATCATGATTTTGTGATAAAACCTTCGCGAC 1713
Db |||||
Qy 1993 AACATAACAGACAAATTTGATCAGATTAATCATGATTTTGTGATAAAACCTTCGCGAC 2052
Db |||||
Qy 1714 CAGGGGGCAATGCAATTTGGTGGACAGATGGAGCAATGGATACCGCGAGGATTTGGA 1773
Db |||||
Qy 2053 CAGGGGGCAATGCAATTTGGTGGACAGATGGAGCAATGGATACCGCGAGGATTTGGA 2112
Db |||||

Qy 1774 GTTACAGGCGTTAATAATTGCGAGTTATCGCTTTATCTGTATATGCAAAATTTGCTTTTAG 1833
Db |||||
Qy 2113 GTTACAGGCGTTAATAATTGCGAGTTATCGCTTTATCTGTATATGCAAAATTTGCTTTTAG 2172
Db |||||
RESULT 6
US-09-949-016-13845/c
; Sequence 13845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13845
; LENGTH: 152132
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13845

Query Match 4.1%; Score 74.6; DB 4; Length 152132;
Best Local Similarity 50.6%; Pred. No. 7.8e-12;
Matches 206; Conservative 0; Mismatches 199; Indels 2; Gaps 1;

Qy 872 CATCAAAATGCCACCCACTCCCTCACCACAAACCAACGACGCCACACAGGAGGAACA 931
Db |||||
Qy 136819 CATCACCACCCACCCACCCACCCACCCACCCCTCCACCCACCCACCCATCACCAC 136760
Db |||||
Qy 932 ACACAAACCATTTCCAAAGATGCTGTGACTGAACCTAGACAAATAACACAACTGCACAAC 991
Db |||||
Qy 136759 CACCACACACACACACACCCACCCACCTCCACCCACCCACCCATCCACCCACCCAC 136700
Db |||||
Qy 992 CGT--COATGCCCTCATTAACACTACCAATCTCTACTAACAACTCTCAAAACAA 1049
Db |||||
Qy 136699 CATCACTATCACCACCCACCCACCCACCCACCCACCCATCACCATCACCACCCAC 136640
Db |||||
Qy 1050 CTTGAGACTCTCTCTGACGATTTAGAAACACACCAACCAATGACAAACAGACACAAT 1109
Db |||||
Qy 136639 CACCACACCCACCCACCCACCCACCCCTCCACCCACCCACCCATCACCACCCACCCAT 136580
Db |||||
Qy 1110 CACTGAAAATGAGCAAAACACGAGTGCCTCGATAACAACTCTCCACCGGAAATCC 1169
Db |||||
Qy 136579 CACACACACACACCTTCCACCCACCAACCATCATCACCACCATCACCACCAATAC 136520
Db |||||
Qy 1170 CACCACAGCAAGAGCAGCAGCAAAAAGGCCCGCCACCAACGGCAGCAACACAGAC 1229
Db |||||
Qy 136519 CACCACACCCACCCACCCACCCACCCACCCACCCACCCATCACCATCACCACCCAC 136460
Db |||||
Qy 1230 AATGAGCATTTTACAGAGTCTCTCCCGCCACCCAGCTCGACTGCAC 1276
Db |||||
Qy 136459 CACCATCACCACCAATATCACCACCCACCCACCCACCCACCCACCCACCC 136413
Db |||||
RESULT 7
US-09-949-016-12371/c
; Sequence 12371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12371
; LENGTH: 152145
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12371

Query Match
Best Local Similarity 4.1%; Score 74.6; DB 4; Length 152145;
Matches 206; Conservative 0; Mismatches 199; Indels 2; Gaps 1;

QY 872 CATCAACATGCGCCCACTCCCTCACCACCAACCAAGCAGCGCCACGAGGAAGGAAACA 931
Db 136819 CATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136760

QY 932 ACACAAACATCCCAAGATGCTGTGATGAATAGACAAATAATACCACTGCACAAAC 991
Db 136759 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136700

QY 992 CGT--CCATGCCCTCTATAACACTACCAATCTCTACTACCAACCTCCCAACAA 1049
Db 136699 CATCACCATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136640

QY 1050 CTTGAGCACTCTCTGCAACCAATTAACAAACCAACCAACCAACCAACCAACCAAC 1109
Db 136639 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136580

QY 1110 CACTGAAATAGCAAAACCAAGTCCCTTCGATACAAACCTGCTCCAAACGGGAAATCC 1169
Db 136579 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136520

QY 1170 CACCACGAAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1229
Db 136519 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136460

QY 1230 AAATGAGCATTTTCCACAGTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1276
Db 136459 CACCATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136413

RESULT 8
US-09-949-016-12052
; Sequence 12052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12052
; LENGTH: 34230
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12052

Query Match
Best Local Similarity 3.3%; Score 60.4; DB 4; Length 128470;
Matches 136; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 1009 AACACTACCAACATCTCTACTAACAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCA 1068
Db 98325 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98384

QY 1059 CCATTACAAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1128
Db 98385 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98444

QY 1129 AGTGCCCTCTGATAAACAACCTTGCCTCCAAACGGGAAATCCCAACCAAGAGAGCACC 1188
Db 98445 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98504

QY 1189 AGCAGCAAAAAGGCCCGCCGACCAACCGGACCAACCAAGCAACCAATGAGCAATTTACCACT 1248
Db 98505 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98564

QY 1249 CCTCCCCCCCCCAGCTCGA 1270
```

```
; Sequence 13765, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13765
; LENGTH: 128470
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13765

Query Match
Best Local Similarity 3.3%; Score 60.4; DB 4; Length 128470;
Matches 136; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 1009 AACACTACCAACATCTCTACTAACAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCA 1068
Db 98325 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98384

QY 1059 CCATTACAAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1128
Db 98385 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98444

QY 1129 AGTGCCCTCTGATAAACAACCTTGCCTCCAAACGGGAAATCCCAACCAAGAGAGCACC 1188
Db 98445 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98504

QY 1189 AGCAGCAAAAAGGCCCGCCGACCAACCGGACCAACCAAGCAACCAATGAGCAATTTACCACT 1248
Db 98505 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98564

QY 1249 CCTCCCCCCCCCAGCTCGA 1270
```

Db 98565 ACCACCACCAACCATCACCA 98586

RESULT 10

US-08-928-361B-2

; Sequence 2, Application US/08928361B

; Patent No. 6071518

; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn

; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

; TITLE OF INVENTION: SPECIES INFECTIONS

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

; STREET: 385 Sherman Avenue, Suite 6

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-1840

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,361B

; FILING DATE: 12-SEP-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026,062

; FILING DATE: 13-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: VERNY, Hana

; REGISTRATION NUMBER: 30,518

; REFERENCE/DOCKET NUMBER: 480.76-1(HV)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-1677

; TELEFAX: 650-324-1678

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5511 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-928-361B-2

Query Match 3.1%; Score 57.6; DB 3; Length 5511;

Best Local Similarity 45.7%; Pred. No. 2.8e-07;

Matches 201; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

```
QY 832 CCCCAACAACTTCTGATCGGTCCACGAGCGGCTTTTCATCAACAATGCCCCACT 891
Db 895 CCATACACTAAATGTTGGAGTGAACACACAAACAACTACTACTACTACT 954
QY 892 CCCTCACCACCAACGACGCGCACGAGGAGGAAACACAAACCAATTTCCCAAGAT 951
Db 955 ACTACTAGCACAACAAACACACGACACAAACAACTACTACTACTACTACT 1014
QY 952 GCTGTGACTGAAGTACAAAAATAACAACTGCAACACCGTCCATGCCCTCATAC 1011
Db 1015 ACTACTAGCACAACAACTACTACTAGCACAACAACTACTACTACTACTACT 1074
QY 1012 ACTTACCAATCTTACTTAACACACCTCCAAACAACTTCAGCACTCTCTGCACA 1071
Db 1075 ACAACCAACAACTACTAACCAACGACACTAACCAACCACTACCAAGAAACCA 1134
QY 1072 TTACAAAAACCAACCAATGACAAACACAGAGACAAATCACTGAAATGAGCAACCCAGT 1131
Db 1135 ACAACCAACAACTACTTAACCAACGACACTTACTACTACTACTACTACTACT 1194
QY 1132 GCCCCTCGATAACCAACCTGCTCCAAACGGGAAATCCCAACAGAAAGACCAACC 1191
Db 1195 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1254
```



```
QY 1132 GCCCCTCGATAACACCTGCTCCAAACGGGAATCCACACAGCAAGAGCAGCAGC 1191
Db 2863 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 2922
QY 1192 AGCAAAAAGGCGCCGACACAGCGCACAAACAGCAAAATGACATTTTCAACGAGTCCT 1251
Db 2923 ACAACTACCAAGAAACCAACACACAAACAAACAAACAAACAACTACTACTACAC 2982
QY 1252 CCCCCCAGCCCGAGTCGAC 1271
Db 2983 ACGACAACAACAACACGAC 3002

RESULT 14
US-09-949-016-14613/c
; Sequence 14613, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14613
; LENGTH: 30656
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30656)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14613

Query Match 2.9%; Score 54.2; DB 4; Length 30656;
Best Local Similarity 14.5%; Pred. No. 1.3e-05;
Matches 164; Conservative 0; Mismatches 966; Indels 0; Gaps 0;

QY 142 CATCTGATGGGATTACACTGAGTGGCAAAAGTTGCTGATTCCTTTTGAGGACATCC 201
Db 4633 CAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4574
QY 202 AAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCAAAGAAATGTTGAGTACACAGAGGGGAG 261
Db 4573 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4514
QY 262 GAAGCAAAACATGCTACATATAGTGTACGGATCCCTCTGGAAAAATCCTTGCTGTGTA 321
Db 4513 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4454
QY 322 GATCTCTTACCAACATCGTCTGACTATCCGAAATGCAAAACTATCCATCATATTCAAGT 381
Db 4453 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4394
QY 382 CAAAAACCTTCATGCACAGGGGATCGCCCTTCATTATGGGGAGCATTTTTCGTGTATGAT 441
Db 4393 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4334
QY 442 CGATTGCTCCACAACAATGTACGAGCAAGTCTTCACTGAAGGGGAACATAGCAGCT 501
Db 4333 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4274
QY 502 ATGATTGTCAATAAGACAGTGCACAAATGATTTTCTCGCGCAAGGACAGGATACCGT 561
Db 4273 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4214
```

```
QY 562 CATATGAATCTGACTTCTACTATAATAATTGGACAAGTAGTAACGGAACGCAACGAAT 621
Db 4213 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4154
QY 622 GACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATTTCTACAAAGAACCAATGTGCT 681
Db 4153 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4094
QY 682 CCGTCCAAAAATACCTCCACCACTGCCACAGCCCGTCCGGAGATCAAACTCACAAGCACC 741
Db 4093 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4034
QY 742 CCAACTGATGCCACCAAACTCAATACACGCAACCAAGCAGTGTGATGAGGACCTCGCA 801
Db 4033 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3974
QY 802 ACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACCACTTCTGTATGGTGTACCAAG 861
Db 3973 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3914
QY 862 CAAGGGCTTTTCATCAACAATGCCACCACTCCCTCCACCAACCAAGCAGCAGCA 921
Db 3913 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3854
QY 922 GGAGGAAACAAACAAACCAATCCCAAGATGCTGTGACTGAACTAGACAAAATAACACA 981
Db 3853 NNNNNNAACCACTACCAACCAACCAATCCCAAGATGCTGTGACTGAACTAGACAAAATAACACA 3794
QY 982 ACTGCACAAACGCTCCATGCGCCCTCATACCACTACCACTCTCTACTAACCAACCTCC 1041
Db 3793 ATACAGGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3734
QY 1042 AAACACAACTTTCAGCACTCTCTCTGACCACTTACAAACCAACCAACCAACCAACCA 1101
Db 3733 ACGACCATTTACCATCATCTTACCACTTACCACTTACCACTTACCACTTACCACTTACCA 3674
QY 1102 AGCAATCACTGAAATGAGCAAAACGAGTCCCTCGATGAAACCAACCAACCAACCAACG 1161
Db 3673 ACCACCACTTACCACTTACCACTTACCACTTACCACTTACCACTTACCACTTACCACT 3614
QY 1162 GGAATCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1221
Db 3613 ACCATCAACCAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3554
QY 1222 AACACGCAAAATGAGCAATTTTCCACCACTCTCTCCCTCCCTCCCTCCCTCCCTCCGAC 1271
Db 3553 ATCACCACCACTTACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 3504
```

RESULT 15

```
US-09-949-016-103758/c
; Sequence 103758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103758
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
```

US-09-949-016-103758

```
Query Match      2.9%; Score 54; DB 4; Length 522;
Best Local Similarity 48.7%; Pred. No. 8e-07;
Matches 147; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 969 CAAAATAACAACTGGCACAACCGTCCATATGCCCTCATTAACACTTACCAATCTCTAC 1028
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
487 CATCACCACCACCACAATCACCAGCACCACCAACACCAACCAACCAACCAAT 428

QY 1029 TAACAACACCTCCAACACAACCTTCAGCACTCTCTGCAACCATTAACAAACACCAAC 1088
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
427 CACCAGCACCCCATCACCACCACTTACCATCTACTACCATCAATCAATCACCAGCAC 368

QY 1089 TGACAACACACAGAGCACAAATCACTGMAAATGAGCAACCCAGTGCCCTCGATAACAAC 1148
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 CAGACCAACACCATCAGAACCACTACCACTACCATCACTACCATCACTACCATCACTAC 308

QY 1149 CCTGCCTCCAACGGGAAATCCCAACACAGCAAGAGCACCAGCAAAAGGCCCCGC 1208
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 CATTACSAACCATCACTTACCACCATCACCACAATCACCAGCACCAGCACCACCAACCAT 248

QY 1209 CACAAGGCACCAACACAGCAAAATGAGCATTTTACCACTCTCCCCCACCAGCTC 1268
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 CACAACCACTTACCACCATCACCACCACTTACCATCATCATCAATCAATCACCAGCAC 188

QY 1269 GA 1270
Db 187 CA 186
```

Search completed: August 1, 2005, 10:31:08
Job time : 298.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1157.01 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgagaccacatgttctct.....ttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1841	100.0	1841	15	US-10-066-506A-3
2	1308	71.0	2046	15	US-10-066-506A-5
3	1303.4	70.8	7005	20	US-10-491-121-31
4	1303.4	70.8	7778	20	US-10-491-121-30
5	1303.4	70.8	8256	20	US-10-491-121-32
6	1303.4	70.8	11460	16	US-10-267-322-1
7	1274	69.2	2051	15	US-10-066-506A-9

8	1148.2	62.4	2046	21	US-10-811-353-2	Sequence 2, Appli
9	1148.2	62.4	19112	17	US-10-353-856-27	Sequence 27, Appl
10	780	42.4	2046	15	US-10-066-506A-7	Sequence 7, Appli
11	780	42.4	2046	15	US-10-066-506A-13	Sequence 13, Appl
12	688	37.4	6902	20	US-10-491-121-42	Sequence 42, Appl
13	540	29.3	2039	15	US-10-066-506A-11	Sequence 11, Appl
14	527.6	28.7	7272	22	US-10-860-878-4	Sequence 4, Appli
15	527.6	28.7	7285	22	US-10-860-878-3	Sequence 3, Appli
16	527.6	28.7	18959	17	US-10-353-856-19	Sequence 19, Appl
17	527.6	28.7	18959	17	US-10-353-856-37	Sequence 37, Appl
18	527.6	28.7	18959	17	US-10-353-856-46	Sequence 46, Appl
19	526	28.6	2298	9	US-09-337-946A-1	Sequence 1, Appli
20	526	28.6	2298	17	US-10-384-976-1	Sequence 1, Appli
21	526	28.6	2298	18	US-10-226-795-1	Sequence 1, Appli
22	526	28.6	2298	22	US-10-696-633-1	Sequence 10, Appl
23	526	28.6	6467	20	US-10-491-121-10	Sequence 3, Appli
24	526	28.6	6624	20	US-10-491-121-3	Sequence 2, Appli
25	526	28.6	7154	20	US-10-491-121-1	Sequence 2, Appli
26	526	28.6	7188	20	US-10-491-121-2	Sequence 28, Appl
27	526	28.6	8439	20	US-10-491-121-28	Sequence 4, Appli
28	525.6	28.5	6561	20	US-10-491-121-4	Sequence 4, Appli
29	514.4	27.9	2030	21	US-10-811-353-4	Sequence 9, Appli
30	451.8	24.5	6914	20	US-10-491-121-9	Sequence 29, Appl
31	451.8	24.5	8199	20	US-10-491-121-29	Sequence 25, Appl
32	451.8	24.5	10783	20	US-10-491-121-25	Sequence 7, Appli
33	420	22.8	7106	20	US-10-491-121-8	Sequence 8, Appli
34	336.2	18.3	7044	20	US-10-491-121-7	Sequence 16, Appl
35	324.6	17.6	7002	20	US-10-491-121-16	Sequence 21, Appl
36	324.6	17.6	7023	20	US-10-491-121-21	Sequence 17, Appl
37	324.6	17.6	7036	20	US-10-491-121-17	Sequence 1, Appli
38	324.2	17.6	7003	22	US-10-860-878-1	Sequence 37, Appl
39	290.4	15.8	6324	20	US-10-491-121-37	Sequence 12, Appl
40	286.6	15.6	2360	17	US-10-397-635-12	Sequence 19, Appl
41	285	15.5	6889	20	US-10-491-121-19	Sequence 20, Appl
42	285	15.5	8146	20	US-10-491-121-20	Sequence 9, Appli
43	283.4	15.4	18890	17	US-10-353-856-9	Sequence 2, Appli
44	271.6	14.8	7073	22	US-10-860-878-2	Sequence 13, Appl
45	271.6	14.8	7082	20	US-10-491-121-13	

ALIGNMENTS

RESULT 1

US-10-066-506A-3
; Sequence 3, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 2
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506A-3

Query Match 100.0%; Score 1841; DB 15; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGACCATGTTTCCTTATCATCTTATCTATCAAGGACAAAATCTCCC 60

; OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 71.0%; Score 1308; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0;

Qy	1	ATGAAGACCA	CATGTTTCCTTTATCAGTCTTATCTTAATTCAAGGGACAAAAATCTCCCC	60
Db	1	ATGAAGACCA	CATGTTTCCTTTATCAGTCTTATCTTAATTCAAGGGACAAAAATCTCCCC	60
Qy	61	ATTTTAGAGAT	AGCTAGTAAATAATCAACCCCAAATGTGGATTGCGTATGCTCGGGAAC	120
Db	61	ATTTTAGGAT	AGCTAGTAAATAATCAACCCCAAATGTGGATTGCGTATGCTCGGGAAC	120
Qy	121	CTCAGAAGACA	GAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	180
Db	121	CTCAGAAGACA	GAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	180
Qy	181	GAATCCCTTT	GGAGGATCAAGCGATGGGCTTTCAAGACAGGTGTACCTCCCAAGAAT	240
Db	181	GAATCCCTTT	GGAGGCATCCAGGATGGGCTTTCAAGACAGGTGTACCTCCCAAGAAT	240
Qy	241	GTTGAGTACA	CAGAGGGGGAGAACCAAAACATGCTACAATAAAGTGTAAACGATCCC	300
Db	241	GTTGAGTACA	CAGAGGGGGAGAGCCAAACATGCTACAATAAAGTGTAAACGATCCC	300
Qy	301	TCCTGGAAAT	TCCTGTAGATCCCTACCAACATCCCGTGACTATCCGAATGCAAA	360
Db	301	TCCTGGAAAT	TCCTGTAGATCCCTACCAACATCCCGTGACTATCCGAATGCAAA	360
Qy	361	ACTATCCAT	CATATCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG	420
Db	361	ACTATCCAT	CATATCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG	420
Qy	421	GGAGCATTTT	TCGTATGATCGATGGCTCCACAACAATGTGTACCGAGGCAAACTCTTC	480
Db	421	GGAGCATTTT	TCGTATGATCGATGGCTCCACAACAATGTGTACCGAGGCAAACTCTTC	480
Qy	481	ACTGAAGGAA	CATAGCAGCTATGATGTGTCATTAAGACAGGTGCACAAATGATTTTCTCG	540
Db	481	ACTGAAGGAA	CATAGCAGCTATGATGTGTCATTAAGACAGGTGCACAAATGATTTTCTCG	540
Qy	541	CGGCAAGGACA	AGGGTACCGTCAATGAAATCTGACTTAATAATAATATTGGACAAGT	600
Db	541	CGGCAAGGACA	AGGGTACCGTCAATGAAATCTGACTTAATAATAATATTGGACAAGT	600
Qy	601	AGTAAACGGA	ACGCAAAAGAAATGACACTGGATGTTTCGGCGCTCTTCAGAGATACAAATCT	660
Db	601	AGTAAACGGA	ACGCAAAAGAAATGACACTGGATGTTTCGGCGCTCTTCAGAGATACAAATCT	660
Qy	661	ACAAAGAAC	CAAAACATGTGCTCCGTCCAAAATACCTCCACCACTGGCCACAGCCGCTCG	720
Db	661	ACAAAGAAC	CAAAACATGTGCTCCGTCCAAAATACCTCCACCACTGGCCACAGCCGCTCG	720
Qy	721	GAGATCAAA	CTCACAGCACCCCAACTGATGCGACCAAACTCAATACCAACGGAACCAAGC	780
Db	721	GAGATCAAA	CTCACAGCACCCCAACTGATGCGACCAAACTCAATACCAACGGAACCAAGC	780
Qy	781	AGTCATGAT	GAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACAGAAACCCACACA	840
Db	781	AGTCATGAT	GAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACAGAAACCCACACA	840
Qy	841	ACTTCTGAT	GCGGTACCAAGCAAGGGCTTTTCATCAACAAATGCGCACCTCCCTCACCA	900
Db	841	ACTTCTGAT	GCGGTACCAAGCAAGGGCTTTTCATCAACAAATGCGCACCTCCCTCACCA	900
Qy	901	CAACCAAGCA	CGCCACGCAAGGAGGAAACAAACAAACCATTTCCCAAGATGCTGTGACT	960
Db	901	CAACCAAGCA	CGCCACGCAAGGAGGAAACAAACAAACCATTTCCCAAGATGCTGTGACT	960
Qy	961	GAACTAGACA	AAAAATAACAACTGTGCACAACCGTCCATGCCCTTCATAAACACTACCACA	1020

Db	961	GAAC	TAGACAA	AAATTA	CACTG	CAACCG	TCCATG	CGCCCT	CTATTA	CACTAC	CA	1021		
Qy	1021	ATCT	CTACTA	CAACAC	CTTCAA	ACACAA	CACTTCT	CTGAC	ACTTCT	CTGCA	CAATTA	CAAAAC	1080	
Db	1021	ATCT	CTACTA	CAACAC	CTTCAA	ACACAA	CACTTCT	CTGAC	ACTTCT	CTGCA	CAATTA	CAAAAC	1080	
Qy	1081	ACC	CAAC	CAATG	ACAC	CAAG	CAACA	CAATCA	CTG	AAATG	AGCA	AAACCA	GTGCGCCCTCG	1140
Db	1081	ACC	CAAC	CAATG	ACAC	CAAG	CAACA	CAATCA	CTG	AAATG	AGCA	AAACCA	GTGCGCCCTCG	1140
Qy	1141	ATA	CAAC	CCCTG	CTTCAA	CGG	GAATTC	CCAC	CA	CGG	GAATTC	CCAC	CA	1200
Db	1141	ATA	CAAC	CCCTG	CTTCAA	CGG	GAATTC	CCAC	CA	CGG	GAATTC	CCAC	CA	1200
Qy	1201	GGC	CCG	CGC	CAAC	AGG	CAAC	CA	CA	CA	CA	CA	CA	1260
Db	1201	GGC	CCG	CGC	CAAC	AGG	CAAC	CA	CA	CA	CA	CA	CA	1260
Qy	1261	CCC	AGTCT	CGACT	GCACA	CAATCT	TGTATATTT	TTC	AGA	AGAA	GC	GATCG	1308	
Db	1261	CCC	AGTCT	CGACT	GCACA	CAATCT	TGTATATTT	TTC	AGA	AGAA	GC	GATCG	1308	

RESULT 3

US-10-491-121-31

Sequence 31, Application US/10491121

Publication No. US20040259825A1

GENERAL INFORMATION:

APPLICANT: NABEL, GARY

APPLICANT: YANG, ZHI-YONG

APPLICANT: SULLIVAN, NANCY

APPLICANT: SANCHEZ, ANTHONY

TITLE OF INVENTION: Development of a Preventive Vaccine for

TITLE OF INVENTION: Filovirus Infection in Primates

FILE REFERENCE: NIH221.001NP

CURRENT APPLICATION NUMBER: US/10/491,121

CURRENT FILING DATE: 2004-03-26

PRIOR APPLICATION NUMBER: PCT/US02/30251

PRIOR FILING DATE: 2002-09-24

PRIOR APPLICATION NUMBER: US 60/326476

PRIOR FILING DATE: 2001-10-01

NUMBER OF SEQ ID NOS: 52

SOFTWARE: Fast-Seq for Windows Version 4.0

SEQ ID NO 31

LENGTH: 7005

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg GP (dTM)

US-10-491-121-31

Query Match	70.8%;	Score 1303.4;	DB 20;	Length 7005;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 1304;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Qy	1	ATGAAGACCA	CGTCTTCC	TATCAAGGGACAAAAATCTCCCC	60
Db	2034	ATGAAGACCA	CGTCTTATCTTAATTC	CAAGGGACAAAAATCTCCCC	2093
Qy	61	ATTTTAGAGATAGCTAGTAA	TATCAACCCCAAAATGTG	GATTCGGTGATGCTCCGGAACT	120
Db	2094	ATTTTAGAGATAGCTAGTAA	TATCAACCCCAAAATGTG	GATTCGGTGATGCTCCGGAACT	2153
Qy	121	CTCCAGAAGACAGAAGAC	GTCCATCTGATGGGATTC	CATGTAGTGGGCAAAAGTTGCT	180
Db	2154	CTCCAGAAGACAGAAGAC	GTCCATCTGATGGGATTC	CATGTAGTGGGCAAAAGTTGCT	2213
Qy	181	GATTTC	CCCTTTTGAGGGCATCA	AGCGATGGCTTTCAGGACAGGTGTACCTCCCAAGAAT	240
Db	2214	GATTTC	CCCTTTTGAGGGCATCA	AGCGATGGCTTTCAGGACAGGTGTACCTCCCAAGAAT	2273
Qy	241	GTTGAGTACACAGAGGGG	GGAAGCAAAACATGCTACA	ATAATAGTGTAAACGGATCCC	300

```
Db 2274 GTTGAGTACACAGAGGGGAGGAAAGCCAAACATGCTACAATATAGTGTAAAGGATCC 2333
QY 301 TCTGGAAATCTTGCTGTAGATCTCTCAACAACATCCGTGACTATCCGAAATGCAAA 360
Db 2334 TCTGGAAATCTTGCTGTAGATCTCTCAACAACATCCGTGACTATCCGAAATGCAAA 2393
QY 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG 420
Db 2394 ACTATCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG 2453
QY 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACAATGTACGGAGGCAAGTCTTC 480
Db 2454 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACAATGTACGGAGGCAAGTCTTC 2513
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 540
Db 2514 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 2573
QY 541 CGGCAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTGGAAGA 600
Db 2574 CGGCAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTGGAAGA 2633
QY 601 AGTAACGGAACGCAACAGATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 660
Db 2634 AGTAACGGAACGCAACAGATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 2693
QY 661 ACAAGAACCAACATGTCCTCGTCCAAATACCTCCACACTGCCACAGCCCGTCG 720
Db 2694 ACAAGAACCAACATGTCCTCGTCCAAATACCTCCACACTGCCACAGCCCGTCG 2753
QY 721 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 780
Db 2754 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 2813
QY 781 AGTGATGATGAGACCTCCCAACATCCGCTCAGGTCGGAGAGAGGAAACCCACACA 840
Db 2814 AGTGATGATGAGACCTCCCAACATCCGCTCAGGTCGGAGAGAGGAAACCCACACA 2873
QY 841 ACTTCTGATCGGTACCAAGCAAGGCTTTTATCAACAATGCCACCACTCCCTCACC 900
Db 2874 ACTTCTGATCGGTACCAAGCAAGGCTTTTATCAACAATGCCACCACTCCCTCACC 2933
QY 901 CAACCAAGCACGCCACAGAGGAGGAAACAACAACCATTTCCCAAGATCTGTGACT 960
Db 2934 CAACCAAGCACGCCACAGAGGAGGAAACAACAACCATTTCCCAAGATCTGTGACT 2993
QY 961 GAACCTAGACAAAATTAACAACAACCTGCAACACCGTCCATGCCCTCATTAACACTACC 1020
Db 2994 GAACCTAGACAAAATTAACAACAACCTGCAACACCGTCCATGCCCTCATTAACACTACC 3053
QY 1021 ATCTCTACTAAACAACCTCCAAACAACACTTCAGCACTCTCTGCAACCATTTACAAAC 1080
Db 3054 ATCTCTACTAAACAACCTCCAAACAACACTTCAGCACTCTCTGCAACCATTTACAAAC 3113
QY 1081 ACCACCAATGACAAACACAGAGGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG 1140
Db 3114 ACCACCAATGACAAACACAGAGGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG 3173
QY 1141 ATAAACAACCTGCTCCAAAGGGAATCCCAACAGAAAGAGACAGAGGAGGAAAGAA 1200
Db 3174 ATAAACAACCTGCTCCAAAGGGAATCCCAACAGAAAGAGACAGAGGAGGAAAGAA 3233
QY 1201 GGCCCGCCACAGGACCAACAGCAAAATGACATTTTCCACAGTCCCTCCCCAC 1260
Db 3234 GGCCCGCCACAGGACCAACAGCAAAATGACATTTTCCACAGTCCCTCCCCAC 3293
QY 1261 CCCAGCTCGACTGCAACACATCTTGATATTTTTCAGAAAGGCGA 1305
Db 3294 CCCAGCTCGACTGCAACACATCTTGATATTTTTCAGAAAGGCGA 3338
```

RESULT 4
US-10-491-121-30

```
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30
```

Query Match 70.8%; Score 1303.4; DB 20; Length 7778;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAAGACCACATGTTCTCTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCCC 60
Db 2034 ATGAAGACCACATGTTCTCTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCCC 2093
QY 61 ATTTAGAGTAGTAGTAATAAATCAACCCCAAAATGGATTCCGTATGTCCTCGACT 120
Db 2094 ATTTAGAGTAGTAGTAATAAATCAACCCCAAAATGGATTCCGTATGTCCTCGACT 2153
QY 121 CTCAGAGACAGAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT 180
Db 2154 CTCAGAGACAGAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT 2213
QY 181 GATTCCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTCAAGACAGGTGTACTCCCAAGAA 240
Db 2214 GATTCCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTCAAGACAGGTGTACTCCCAAGAA 2273
QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACNATATAGTGTAAAGTAAAGGATCC 300
Db 2274 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACNATATAGTGTAAAGTAAAGGATCC 2333
QY 301 TCTGAAAATCTCTGCTGTAGATCTCTTACCAACATCCCGTACTATCCGAAATGCAAA 360
Db 2334 TCTGAAAATCTCTGCTGTAGATCTCTTACCAACATCCCGTACTATCTTAATGCAAA 2393
QY 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG 420
Db 2394 ACTATCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG 2453
QY 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACAATGTACGGAGGCAAGTCTTC 480
Db 2454 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACAATGTACGGAGGCAAGTCTTC 2513
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 540
Db 2514 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 2573
QY 541 CGGCAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTGGAAGA 600
Db 2574 CGGCAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTGGAAGA 2633
QY 601 AGTAACGGAACGCAACAGATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 660
Db 2634 AGTAACGGAACGCAACAGATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 2693
```

```
QY 661 ACAAGAACCAAAACATGTGCTCGTCCAAATACTCTCCACCACTGCGCCACAGCCCGTCCG 720
Db |||||||
QY 2694 ACAAGAACCAAAACATGTGCTCGTCCAAATACTCTCCACCACTGCGCCACAGCCCGTCCG 2753
Db |||||||
QY 721 GAGATCAAACTCAAGACACCCCAACTGATGCCCAAACTCAATACCAAGACCCCAAGC 780
Db |||||||
QY 2754 GAGATCAAACTCAAGACACCCCAACTGATGCCCAAACTCAATACCAAGACCCCAAGC 2813
Db |||||||
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 840
Db |||||||
QY 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 2873
Db |||||||
QY 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCAATCAATGCCACCCCACTCCCTCACA 900
Db |||||||
QY 2874 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCAATCAACAAATGCCACCCCACTCCCTCACA 2933
Db |||||||
QY 901 CAACCAAGCAGCCCAACAGCAAGGAGAAACACACAAACCAATTTCCCAAGATGCTGTGACT 960
Db |||||||
QY 2934 CAACCAAGCAGCCCAACAGCAAGGAGAAACACAAACCAATTTCCCAAGATGCTGTGACT 2993
Db |||||||
QY 961 GAACTAGACAAAATAAACAACACTGTCACAAACCGTCCATGCCGCCCTCATAAACAACCTACCA 1020
Db |||||||
QY 2994 GAACTAGACAAAATAAACAACACTGTCACAAACCGTCCATGCCGCCCTCATAAACAACCTACCA 3053
Db |||||||
QY 1021 ATCTCTACTAACAACACTCTCCAAACACAACTTTGAGGACCTCTCTCTGACCAATTAACAAAC 1080
Db |||||||
QY 3054 ATCTCTACTAACAACACTCTCCAAACACAACTTTGAGGACCTCTCTCTGACCAATTAACAAAC 3113
Db |||||||
QY 1081 ACCACCAATGACACACACAGACGACAACTCACTGAAATGAGCAACACAGTGCCGCCCTCG 1140
Db |||||||
QY 3114 ACCACCAATGACACACACAGACGACAACTCACTGAAATGAGCAACACAGTGCCGCCCTCG 3173
Db |||||||
QY 1141 ATAAACAACCTCGCTCCAAACGGGAATCCCAACACAGCAAGAGACACAGCAAGCAAAAAA 1200
Db |||||||
QY 3174 ATAAACAACCTCGCTCCAAACGGGAATCCCAACACAGCAAGAGACACAGCAAGCAAAAAA 3233
Db |||||||
QY 1201 GSCCGCGGCAACAGCGCACCAAGCAAGCAATGAGCAATTTCAACAGTCTCTCCGCCCAACC 1260
Db |||||||
QY 3234 GSCCGCGGCAACAGCGCACCAAGCAAGCAATGAGCAATTTCAACAGTCTCTCCGCCCAACC 3293
Db |||||||
QY 1261 CCAGCTGACCTGCACACATCTCTATATTTTCAAGAAAGCGA 1305
Db |||||||
QY 3294 CCAGCTGACCTGCACACATCTCTATATTTTCAAGAAAGCGA 3338
Db |||||||
```

RESULT 5

```
US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication NO. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 8256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Marburg GP (dTM)
US-10-491-121-32
```

```
Query Match 70.8%; Score 1303.4; DB 20; Length 8256;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACCAACATCTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAAATCTCCCC 60
Db |||||||
QY 1431 ATGAGACCAACATCTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAAATCTCCCC 1490
Db |||||||
QY 61 ATTTTAGAGATAGCTAGTAATAAACAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120
Db |||||||
QY 1491 ATTTTAGAGATAGCTAGTAATAAACAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 1550
Db |||||||
QY 121 CTCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAAGTTGCT 180
Db |||||||
QY 1551 CTCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAAGTTGCT 1610
Db |||||||
QY 181 GATTCCCTTTGGAGGCATCCAGGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db |||||||
QY 1611 GATTCCCTTTGGAGGCATCCAGGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 1670
Db |||||||
QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATAATAAGTGTAAACGGATCCC 300
Db |||||||
QY 1671 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATAATAAGTGTAAACGGATCCC 1730
Db |||||||
QY 301 TCTGGAAAAATCCTTGTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db |||||||
QY 1731 TCTGGAAAAATCCTTGTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 1790
Db |||||||
QY 361 ACTATCCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATGCGCTTCATTATGG 420
Db |||||||
QY 1791 ACTATCCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATGCGCTTCATTATGG 1850
Db |||||||
QY 421 GGAGCATTTTTCTGTATGATGCGCATTTGCCCTCCACAACATGTACCGAGGCAAAAGTCTTC 480
Db |||||||
QY 1851 GGAGCATTTTTCTGTATGATGCGCATTTGCCCTCCACAACATGTACCGAGGCAAAAGTCTTC 1910
Db |||||||
QY 481 ACTGAAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCG 540
Db |||||||
QY 1911 ACTGAAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCG 1970
Db |||||||
QY 541 CGGCAAGGACAGAGGTACCGTCAATGATCTGATCTTCTACTAATAATAATTTGGACAAGT 600
Db |||||||
QY 1971 CGGCAAGGACAGAGGTACCGTCAATGATCTGATCTTCTACTAATAATAATTTGGACAAGT 2030
Db |||||||
QY 601 AGTAAACGGAACGCAAAACGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTC 660
Db |||||||
QY 2031 AGTAAACGGAACGCAAAACGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTC 2090
Db |||||||
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCACTGCCCCACAGCCCGTCCG 720
Db |||||||
QY 2091 ACAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCACTGCCCCACAGCCCGTCCG 2150
Db |||||||
QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
Db |||||||
QY 2151 GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 2210
Db |||||||
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 840
Db |||||||
QY 2211 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 2270
Db |||||||
QY 841 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCAATCAACAAATGCCACCCCACTCCCTCACA 900
Db |||||||
QY 2271 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCAATCAACAAATGCCACCCCACTCCCTCACA 2330
Db |||||||
QY 901 CAACCAAGCAGCCCAACAGCAAGGAGGAAACACAAACCAATTTCCCAAGATGCTGTGACT 960
Db |||||||
QY 2331 CAACCAAGCAGCCCAACAGCAAGGAGGAAACACAAACCAATTTCCCAAGATGCTGTGACT 2390
Db |||||||
QY 961 GAACTAGACAAAATAAACAACACTGCAACCGTCCATGCCGCCCTCATAAACAACCTACCA 1020
Db |||||||
QY 2391 GAACTAGACAAAATAAACAACACTGCAACCGTCCATGCCGCCCTCATAAACAACCTACCA 2450
Db |||||||
```


APPLICANT: Hevey, Michael C.
APPLICANT: Schmaljohn, Alan, L.
TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
FILE REFERENCE: 003/243/SAP
CURRENT APPLICATION NUMBER: US/10/066,506A
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/267,522
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 9
LENGTH: 2051
TYPE: DNA
ORGANISM: Marburg virus strain Musoke
FEATURE:
OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg virus Glycoprotein 2
OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-9

Query Match 69.2%; Score 1274; DB 15; Length 2051;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 60
DB 10 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 69
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAAATGTGGATTGCTTCGGAACT 120
DB 70 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAAATGTGGATTGCTTCGGAACT 129
QY 121 CTCGAGAGACAGAGACGTCATCTGATGGGATTCACTGAGTGGGCAAAAATGTTGCT 180
DB 130 CTCGAGAGACAGAGACGTCATCTGATGGGATTCACTGAGTGGGCAAAAATGTTGCT 189
QY 181 GATTCCTCTTGGAGGATCCAAAGGATCGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
DB 190 GATTCCTCTTGGAGGATCCAAAGGATCGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 249
QY 241 GTTGGATCACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC 300
DB 250 GTTGGATCACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC 309
QY 301 TCTGGAATAATCTTGTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
DB 310 TCTGGAATAATCTTGTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 369
QY 361 ACTATCCATCATTAATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 420
DB 370 ACTATCCATCATTAATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 429
QY 421 GGAGCATTTTTTCTGTATGATCGCATTTGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
DB 430 GGAGCATTTTTTCTGTATGATCGCATTTGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 477
QY 481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAGACAGTGCACAAAATGATTTTCG 540
DB 478 ACTGAAGGAAACATAGCAGCTATGATGTCAATAGACAGTGCACAAAATGATTTTCG 537
QY 541 CGGACAGGACNAGGGTACCGTCATATGATCTGACTTCTACTAATAATATTTGGACNAGT 600
DB 538 CGGACAGGACNAGGGTACCGTCATATGATCTGACTTCTACTAATAATATTTGGACNAGT 597
QY 601 AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT 660
DB 598 AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT 657
QY 661 ACAAGAACCAAAACATGTGCTCGTCCAAAATACCTCCACCACTGCCACAGCCCGTCG 720
DB 658 ACAAGAACCAAAACATGTGCTCGTCCAAAATACCTCCACCACTGCCACAGCCCGTCG 717
QY 721 GAGATCAAACTCAAGACACCCCACTGATGCGACCAAACTCAATACACAGGACCCCAAGC 780

DB 718 GAGATCAAACTCAAGACACCCCACTGATGCGACCAAACTCAATACACGAGCCCAAGC 777
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACA 840
DB 778 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACA 837
QY 841 ACTTCTGATGCGGTCCCAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACC 900
DB 838 ACTTCTGATGCGGTCCCAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACC 897
QY 901 CAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT 960
DB 898 CAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT 957
QY 961 GAATAGACAAAATAACACAACTGCAACACCGTGCATGCCCTCTATAACACTACCA 1020
DB 958 GAATAGACAAAATAACACAACTGCAACACCGTGCATGCCCTCTATAACACTACCA 1017
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1080
DB 1018 ATCTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1077
QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCTCG 1140
DB 1078 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCTCG 1137
QY 1141 ATAAACACCTCTCCAAACGGGAAATCCCAACACAGCAAGAGACAGCAGCAAGAAAA 1200
DB 1138 ATAAACACCTCTCCAAACGGGAAATCCCAACACAGCAAGAGACAGCAGCAAGAAAA 1197
QY 1201 GGCCCCGCCCAACGGCAGCAACACAGCAAAATGAGCAATTCACCACTCTCCCTCCCAACC 1260
DB 1198 GGCCCCGCCCAACGGCAGCAACACAGCAAAATGAGCAATTCACCACTCTCCCTCCCAACC 1257
QY 1261 CCCAGCTCGACTGCAACCAACATCTTGTATATTTTCAGAAAGAACGATCG 1308
DB 1258 CCCAGCTCGACTGCAACCAACATCTTGTATATTTTCAGAAAGAACGATCG 1305

RESULT 8
US-10-811-353-2
; Sequence 2, Application US/10811353
; Publication No. US20050112098A1
; GENERAL INFORMATION:
; APPLICANT: McCray, Paul B.
; APPLICANT: Sanders, David A.
; APPLICANT: Jeffers, Scott A.
; APPLICANT: Davidson, Beverly L.
; APPLICANT: Simm, Patrick L.
; TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE
; FILE REFERENCE: 290.00670120
; CURRENT APPLICATION NUMBER: US/10/811,353
; CURRENT FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Marburg virus glycoprotein
US-10-811-353-2

Query Match 62.4%; Score 1148.2; DB 21; Length 2046;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 60
DB 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAAATGGAATTCGGTATGCTCCGGAAT 120
DB 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAAATGGAATTCGGTATGCTCCGGAAT 120

QY 121 CTCGAGACAGAGACGTCCATCTGATGGGATTCACTGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCGAGACAGAGAGATCCATCTGATGGGATTCACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGCATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAA 240
DB 181 GATTCCCTTTGGAGCATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAA 240
QY 241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATTAAGTGTAAACGGATCCC 300
DB 241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATTAAGTGTAAACGGATCCC 300
QY 301 TCTGMAAATCTTGTGTTAGATCTCTCAACATCCGTTGATCTCCGAAATGCAAA 360
DB 301 TCTGMAAATCTTGTGTTAGATCTCTCAACATCCGTTGATCTCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTCATTTATGG 420
DB 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTCATTTATGG 420
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACCGAGGCAAAAGTCTTC 480
DB 421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACCGAGGCAAGTCTTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAAGACAGATGCAAAAATGATTTCTCG 540
DB 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAAGACAGATGCAAAAATGATTTCTCG 540
QY 541 CGGCAAGGACAGGGTACCGTATGATGATCTGATCTGATCTTCTACTAATAATATTTGGACAGT 600
DB 541 CGGCAAGGACAGGGTACCGTATGATGATCTGATCTGATCTTCTACTAATAATATTTGGACAGT 600
QY 601 AGTAACGGAAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 AGTAACGGAAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 ACAAGAAACCAACATGCTCGTCCAAATACCTTCAACATGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 ACAAGAAACCAACATGCTCGTCCAAATACCTTCAACATGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GAGTCAAACTCACAGACCCCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GAGTCAAACTCACAGACCCCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 AGTGATGATGAGGACCTCGCAACATCCGCTCAGGTCGAGGAGACGAAACCCACACA 840
DB 781 AGTGATGATGAGGACCTCGCAACATCCGCTCAGGTCGAGGAGACGAAACCCCTATACA 840
QY 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAACAATGCCACCCCACTCCCTCACCA 900
DB 841 ACTTCTGATGCGGTCACTAAGCAAGGCTTTTCAACAATGCCACCCCACTCCCTCACCA 900
QY 901 CAACCAAGCACCCCAAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
DB 901 CAACCAAGCACCCCAAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
QY 961 GAACTAGACAAAAATAACAACCTGCAACACCGTCCATGCCCTCATAACTACTACCACA 1020
DB 961 GAACTAGACAAAAATAACAACCTGCAACACCGTCCATGCCCTCATAACTACTACCACA 1020
QY 1021 ATCTCTACTAAACAACCTTCCAAACCAACTTTTCTGACACTCTCTGACCACTTACAAAC 1080
DB 1021 ATCTCTACTAAACAACCTTCCAAACCAACTTTTCTGACACTCTCTGACCACTTACAAAC 1080
QY 1081 ACCACCAATGACACACAGAGGACCAATCACTGAAATGACAAACCAAGTGCCTCCCTCG 1140
DB 1081 ACCACCAATGACACACAGAGGACCAATCACTGAAATGACAAACCAAGTGCCTCCCTCG 1140
QY 1141 ATAACAACCTTCTCCAAAGGAAATCCCAACAGCAAGGACCAAGGACCAAGGACCAAGGAA 1200
DB 1141 ATAACAACCTTCTCCAAAGGAAATCCCAACAGCAAGGACCAAGGACCAAGGACCAAGGAA 1200
QY 1201 GGGCCCGCCCAACAGGCAACAAACAGCAAAATAGAGCATTTTACCAGTCTCTCCCCACC 1260

DB 1201 GGGCCCAACCAAGGCAACCAATATGACAAATGGGCAATTTAACAGTCCCTCCCCACC 1260
QY 1261 CCCAGCTCGATGCAACACATCTTTGTATATTTTTCAGAAAGCAAGCGA 1305
DB 1261 CCCAACCCGACCAACACATCTTTGTATATTTTCAGAAAGNAACGA 1305
RESULT 9
US-10-353-856-27
; Sequence 27, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawaoka, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Base
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; PRIOR FILING DATE: 2003-01-29
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Marburg virus
US-10-353-856-27
Query Match 62.4%; Score 1148.2; DB 17; Length 19112;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 60
DB 5940 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 5999
QY 61 ATTTTAGAGATAGTGTAAATCAACCCCAAAATGGGATTCGGATTCGTCTCCGAACT 120
DB 6000 ATTTTAGAGATAGTGTAAATCAACCCCAAAATGGGATTCGGATTCGTCTCCGAACT 6059
QY 121 CTCGAGACAGAGACGTCCTATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCGAGACAGAGAGATGTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCTTTGGAGGATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAA 240
DB 6120 GATTCCCTTTGGAGGATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAA 6179
QY 241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATTAAGTGTAAACGGATCCC 300
DB 6180 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATTAAGTGTAAACGGATCCC 6239
QY 301 TCTGMAAATCTTGTGTTAGATCTCTCAACATCCGTTGATCTCCGAAATGCAAA 360
DB 6240 TCTGMAAATCTTGTGTTAGATCTCTCAACATCCGTTGATCTCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTAAGGTCAAAAACCTCATGCAAGGGGATCGCCCTCATTTATGG 420
DB 6300 ACTATCCATCATATTAAGGTCAAAAACCTCATGCAAGGGGATCGCCCTCATTTATGG 6359
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACCGAGGCAAGTCTTC 6419
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAAGACAGATGCAAAAATGATTTCTCG 540
DB 6420 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAAGACAGATGCAAAAATGATTTCTCG 6479
QY 541 CGGCAAGGACAGGGTACCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 600

Db 6480 AGGCAAGGACAGGGGTACCGTCACATGAATCTGACTTCTACTATAATAATATTTGGACAGT 6539
Qy 601 AGTAACGGACCAACAGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCTT 660
Db 6540 AACAAATGGAACCAACAGAAATGACACTGGATGTTTTCGGGTCTCTTCAAGAAATACAACTCC 6599
Qy 661 ACAGAGAACCAACATGTGCTCCGTCCTCAAAATACCTCCACCACTGCGCCACAGCCCGTCCG 720
Db 6600 AGAAGAAATCAAAATGTGCTCCGTCCTCAAAATACCTCACCCTGCGCCACAGCCCGTCCA 6659
Qy 721 GAGATCAAACTCAAGACACCCAACTGATGCGCCAACTCAATPACACGACGCCAACG 780
Db 6660 GAGATCAAACTCAAGACACCCAACTGATGCGCCAACTCAATPACACGACGCCAACG 6719
Qy 781 AGTGATGATGAGGACTCGCAACATCCGCTCAGGGTCGGAGAAAGAGAAACCCACACA 840
Db 6720 AATGATGATGAGGACTCATTAATCCGCTCAGGGTCGGAGAAAGAGAAACCCCTATACA 6779
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCAACAATGCCACCCACTCCCTCACCA 900
Db 6780 ACTTCAGATGCGGTCACTAAGCAAGGGCTTTTCAACAATGCCACCCACTCCCTCACCA 6839
Qy 901 CAACCAAGCACGCCACAGCAGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCACGCCACAGCAGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT 6899
Qy 961 GAATAGACAAATAACAACTGCAACAACTGCAACAACTGCAACAACTGCAACAACTGCA 1020
Db 6900 GAACCCAAACAAACAACTGCAACAACTGCAACAACTGCAACAACTGCAACAACTGCA 6959
Qy 1021 ATCTCTACTTAAACAACTGCAACAACTGCAACAACTGCAACAACTGCAACAACTGCA 1080
Db 6960 ATCTCTACTTAAACAACTGCAACAACTGCAACAACTGCAACAACTGCAACAACTGCA 7019
Qy 1081 ACCCAATATGACAAACAGCAGCACTCACTGAAATATGAGCAACCACTGCAACCACTGCA 1140
Db 7020 ACCCAATATGACAAACAGCAGCAGCACTGAAATATGAGCAACCACTGCAACCACTGCA 7079
Qy 1141 ATAAACCTGCTTCCAAACGGGAAATCCACACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1200
Db 7080 AAAACAACTGCTTCCAAACAGGAAATCTTACCAACAGCAAGCAAGCAAGCAAGCAAGCA 7139
Qy 1201 GCGCCGCGCACAAACGGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1260
Db 7140 GCGCCGCGCACAAACGGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 7199
Qy 1261 CCGAGCTGCACTGCAACCACTCTGTATATTTTCAAGAAAGCGA 1305
Db 7200 CCGAAGCGCACCAACCACTCTGTATATTTTCAAGAAAGCGA 7244

RESULT 10

US-10-066-506A-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein

; OTHER INFORMATION: Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-7

Query Match 42.4%; Score 780; DB 15; Length 2046;
Best Local Similarity 74.8%; Pred. No. 3e-232;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

Qy 1 ATGAGACCAATGCTTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAAATCTCCCC 60
Db 1 ATGAGACCAATATATTTCTGATTTAGTCTCAATTTAATCCAAAGTATAAAAAATCTCCCT 60
Qy 61 ATTTTACAGATAGCTAGTAAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGTAGATTGAGTGTGCTCCGGAAC 120
Qy 121 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACCTGAGTGGGCAAAAAATGTGCT 180
Db 121 CTCAGAGACAGAGATGTTTCTGATGGGATTTTACCTGAGTGGGCAAAAAATGTGCT 180
Qy 181 GATTCCCTTTGAGGAGATCCAGCGATGGCTTTTCAAGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCCTTTGAGGAGATCTAAACGATGGCTTTTCAAGACAGGTGTTCCTCCCAAGAA 240
Qy 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACATATAGTGTAAACGATGCC 300
Db 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACATATAGTGTAAACGATGCC 300
Qy 301 TCTGGAAAAATCTTCTGCTTTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAAAAATCTTCTGCTTTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Qy 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAAGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTTATCATATTTCAAGGTCAAAACCTCATGCAAGGGATTCGCCCTTCATTTATGG 420
Qy 421 GGACATTTTTCTGATGATGCTGCTTCCCAACAAATGTAACGAGGCAAGTCTTC 480
Db 421 GGACATTTTTCTGATGATGCTGCTTCCCAACAAATGTAACGAGGCAAGTCTTC 480
Qy 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCT 540
Qy 541 CGGCAAGACAGAGGTACCGTCAATGAACTGCTTCTACTATAATAATATTTGACAGT 600
Db 541 AGGCAAGACAGAGGTATCGTCAATGAACTGCTTCTACTATAATAATATTTGACAGT 600
Qy 601 AGTAACGGACGCAACGAAATGACATGGAATGTTTCCGCGCTCTTCAAGATACAAATCT 660
Db 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTTCGCAATCTCTCCCAAGAAATACA 660
Qy 661 ACAAGAAACCAACATGCTGCTCCGTCCTCAAAATACCTCCACCTGCGCCACAGCCGTCGG 720
Db 661 ACAACCAATCAACATGCTGCTCCGTCCTCAAAATACCTCCACCTGCGCCACAGTAACTCCG 720
Qy 721 GAGATCAAACTCAAGACACCCCAACTGATGCGCCCAAACTCAATACCAAGGACCCCAAGC 780
Db 721 AGCATTCACCTCAAAATACCTCAAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 AGTGATGATGAGGACTCGCAACATCCGCTCAGGGTCGGAGAAAGAGAAACCCACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGCGCTCAGGATCTGGAGAAACAGGGGCCCCACACA 840
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCAACAATGCCACCCACTCCCTCACCA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAAACATGCTCAACATATTTGTCACCTCTCTCACTA 900
Qy 901 CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTTCCCAAGATGCTGTGACT 960
Db 901 CATCCAAGCACCTCAACATGAGCAAAAACAGTACGAATCTCTCCGACATGCTGTAACT 960
Qy 961 GAATAGCAAAAAATAACAACTGCAACACCGCTCCATGCCCCCTCATTAACACTACCA 1020

Db 961 GAGCAAAATGGAACCGACCCAAACAACAACAGCAACGCTCCTCAACAATACTAATAACA 1020
QY 1021 ATCTCTACTAAACAACAACCTTCCAAACAACACTTCAGCACTCTCTGCACTTACAAAC 1080
Db 1021 ACTCCCACTATAACACTCTGAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
QY 1081 ACCACAATGACAACAACAGAGCAACAATCACTGAAAAATGAGCAAAACCAAGTGCCTCCG 1140
Db 1081 ATCAACAATAATGATACACAAGGTGAATAGCAGAAAGCGAAACAACAATGCTCAGTTG 1140
QY 1141 ATAACAACCTCTCCCAACGGGAATCCCAACAGCAAGCAAGAGCAGCAGCAAAAA 1200
Db 1141 AACACAATCTAGATCCCAACAGAAATCCCAACAGAGCAAGCAACCAACCAACCAACC 1200
QY 1201 GCGCCGCGCACAAACGCGCAACCAACAGCAAAATGAGCAATTTCAACAGTCTCCGCCAC 1260
Db 1201 ACATCATCATGAGCATCAGATATACAGCAACCAACCAACCAACCAATTTCTTCGGAT 1260
QY 1261 CCCAGCTCGATGCAACAACATCTTGATATTTTCAAGAAAGCGATCG 1308
Db 1261 TCTAGTCCGACAACCGCCCTCCTATATCTTTAGAAAGAAACGATCG 1308

RESULT 11
US-10-066-506A-13
; Sequence 13, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066, 506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267, 522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 13
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Marburg virus strain Raven
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
US-10-066-506A-13

Query Match 42.4%; Score 780; DB 15; Length 2046;
Best Local Similarity 74.8%; Pred. No. 3e-232;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGCAAAAAATCTCCCC 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAACTCTCCCT 60
QY 61 ATTTTAGACATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTGCTGCTCCGGAAT 120
Db 61 GTTTTAGAAATTTGCTAGTAACGCCAACTCAAGATGTAGATTCAGTGTGCTCCGGAACC 120
QY 121 CTCGAAGACAGAAAGACGCTCCATCTGTGAGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAAAACAGCAAGAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 240
Db 181 GATTCCTCTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCTCCCAAGAAC 240
QY 241 GTTGGATCACAGAGGGGAGGAAGCCAAACATGCTACAATATAAGTGTACAGGATCCC 300
Db 241 GTTGGATATACGGAAGGAAGGAAGCCAAACATGTTTACAAATATAGTGTAAACAGCCCT 300
QY 301 TCTGGAAAATCTCTGCTGTGTAGATCCTCTCAACACATCCGCTGACTATCCGAAATGCAAA 360

Db 301 TCTGGAAAATCTCTGCTGTGTAGATCCTCCAGTAATAATCCGCGATTACCTTAAATGTAAA 360
QY 361 ACTATCCATCATATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCTCATTTATGG 420
Db 361 ACTGTTTCATCATATATTCAAGGTCAAAACCTCATGCACAGGGGATTTGCCCTCATTTTGG 420
QY 421 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAACAATGTATCCGAGGCAAAAGTCTTC 480
Db 421 GGGCATTTTTTCTGTATGATCGCGTTGCCCTCTACAACAATGTATCCGAGGCAAGTCTTTC 480
QY 481 ACTGAAGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAATGATTTTCTTCG 540
Db 481 ACTGAAGGAATATATAGCAGCTATGATTTGTAAATAGACAGTTCACAGAAATGATTTTTTCT 540
QY 541 CGGCAAGCAACAGGTGATCCGTCAATGAACTCTGACTTCTTACTTAATAAATATTGGACAAGT 600
Db 541 AGGCAGAGCAAGGTATTATGTCATAGAACTTGACCTCCACCAATTAATATTGGACAAGC 600
QY 601 AGTAAACGGAACGCAAAACGAATGACATGATGATTTTGGCGCTCTTTCAAGAATACAAATCT 660
Db 601 AGCAATGAACGCGAGAGAAATGATAGGGATGTTTGGCATCTCTCAAGAATACAACTCC 660
QY 661 ACAAGAACAACAACATGCTGCTCCGTCCAAAATACCTCCACCACTGCCCAACAGCCGCTCG 720
Db 661 ACAACAATCAACATGCTGCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 720
QY 721 GAGATCAAACTCACAAGCAGCCCAACTGATGTCACCAAACTCAATACCAAGACCCCAAGC 780
Db 721 AGCATTTCTCTCAAAATACCTCAAAATTAATTAATGCTAAATCTGGAACTATGAAACCAAGT 780
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
QY 841 ACTTCTGATGGGTCAACAAGAGGGCTTTTCATCAACAATGCCCACTCCCTCCCTCACCA 900
Db 841 ACTCTTAATGATGACTCAACAGAAACAATCGTCAACAATATTGTCTCACTCTCTTCACTA 900
QY 901 CAACCAAGCAGCCCAACAGAGGAGGAACAACAACCACTTCCCAAGATGCTGTGACT 960
Db 901 CATCAAGCACTCACAACACTGAGCAAAACAGTACGAATCTCTTCCGACATGCTGTAACT 960
QY 961 GAACTAGACAAAAATAACAACACTGCACAACCGTCCATGCCCCCTCATATAACACTACCA 1020
Db 961 GAGCAATGGAACCGACCCCAACAACAACAACAGCAACGCTCTCTCAACAATACTAATAACA 1020
QY 1021 ATCTCTACTAAACAACCTCCAAACAACAACACTTCAGCACTCTCTCTGCACTTACAAAC 1080
Db 1021 ACTCCCACTATAACACTCTCAAGTACAACTCAGTACTCTCTTCCCTCCCAACCGCAAC 1080
QY 1081 ACCCAATGACAACAACAGAGCAACAATCACTGAAAAATGAGCAAAACAGTGCCTCCCTCG 1140
Db 1081 ATCAACAATATGATACAAACGTTGAACTAGCAGAAAGCAACCAACCAATGCTCAGTTG 1140
QY 1141 ATAACAACCTCTCCCTCCCAACGGGAATCCCAACAGCAAGAGCAAGCAGCAGCAAAAA 1200
Db 1141 AACACAATCTAGATCCCAACAGAAATCCCAACAGGCAAGACACCAACAGCAACAAC 1200
QY 1201 GCGCCGCGCAACAACGGCAACCAACAACAGCAACCAAGATTTTCAACAGTCTCTCCGCCAC 1260
Db 1201 AACATCATCATGAGCATCAGATATACAGATAAACAAGCAACCCCAACAATTTCTTCCGGAT 1260
QY 1261 CCCAGCTCGATGCAACAACATCTTGATATTTTCAAGAAAGCGATCG 1308
Db 1261 TCTAGTCCGACAACCGCCCTCCTATATCTTTAGAAAGAAACGATCG 1308

RESULT 12
US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:

<p> ; APPLICANT: NABEL, GARY ; APPLICANT: YANG, ZHI-YONG ; APPLICANT: SULLIVAN, NANCY ; APPLICANT: SANCHEZ, ANTHONY ; TITLE OF INVENTION: Development of a Preventive Vaccine for ; TITLE OF INVENTION: Filovirus Infection in Primates ; FILE REFERENCE: NIH21.001NP ; CURRENT APPLICATION NUMBER: US/10/491,121 ; CURRENT FILING DATE: 2004-03-26 ; PRIOR APPLICATION NUMBER: PCT/US02/30251 ; PRIOR FILING DATE: 2002-09-24 ; PRIOR APPLICATION NUMBER: US 60/326476 ; PRIOR FILING DATE: 2001-10-01 ; NUMBER OF SEQ ID NOS: 52 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 42 ; LENGTH: 6902 ; TYPE: DNA ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized) ; US-10-491-121-42 </p>									
<p> Query Match 37.4%; Score 688; DB 20; Length 6902; Best Local Similarity 70.5%; Pred. No. 3.7e-203; Matches 919; Conservative 0; Mismatches 385; Indels 0; Gaps 0; </p>									
Qy	1	ATGAAGACCCACATGTTTCCTTATCAGTCTTTATCTTAATTCAAAGGGACAAAAAATCTCCCC	60						
Db	1923	ATGAAGACCCACGCTGCTTTCATCAGCCTGATCTCTGATCCAGGCGCATCAGACCTTGCCC	1982						
Qy	61	ATTTTGAAGATAGCTAGTAATAATCAACCCCAAAATGTGAATTCGGTATGTCCTCGGAAT	120						
Db	1983	ATCTGGAGATCGCCAGCAACAAACAGCCCCAGAAACGTGGACACGCTGTGCAGCGGCACC	2042						
Qy	121	CTCCAGAGACAGAGACGCTCCATCTCATGGGATTTACATGTAGTGGCGCAAAAGTTGCT	180						
Db	2043	CTGCAGAAGACCCGAGGACGTGCACCTGATGGGCTTTCACCTTGAGCGGCCAGAAGGTGCC	2102						
Qy	181	GATTTCCTCTTGAGGAGCATCCAAAGCATGGGCTTTTCAGGACAGAGTGTACCTCCCAAGAT	240						
Db	2103	GACAGCCTCTGGAGGCCAGCAGAGGTGGGCTTTAGGACCGGCTGCCCCCAGAGAC	2162						
Qy	241	GTTCAGTACACAGAGGGGAGGAAGCAAAACATGCTACAATAATAGTGTAAACGGATCCC	300						
Db	2163	GTGCAGTACACCGAGGCGAGGAGCCAAAGACCTGCTACAACATCAGCGTGCAGCGACCCC	2222						
Qy	301	TCTGGAATCTTGTCTGTAGATCTCTTACCAACATCCGTGACTATCCGAATGCAAA	360						
Db	2223	AGCGGCAAGAGCCTGCTGCTGGACCCCTCCCAACCAACATCAGGGACTACCTTAAGTGCAG	2282						
Qy	361	ACTATCCATCATATTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG	420						
Db	2283	ACCATCCACCATCTCAGGGCCAGAACCTCATCGCCACAGGCGATCGCCCTGCACCTGTGG	2342						
Qy	421	GGAGCATTTTTTCTGTATGATCGCATTCGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	480						
Db	2343	GGCGCCTTCTCCTGTACGACAGGATCGCCAGCACCACTGTACAGGGCGAGGGTGTTC	2402						
Qy	481	ACTGAAGGGACATAGCAGCTATGATGTCTAATGAAGACAGTGCACAAAATGATTTCTCG	540						
Db	2403	ACCGAGGGCAACATCGCGGCCATGATCGTTAAAGAACCGGTGCACAGATGATCTTCAGC	2462						
Qy	541	CGGCAGGACAGAGGTACCGTTCATTAATCTGACTTCTACTAATAAATTTGACACAGT	600						
Db	2463	AGGCAGGGCCAGGGCTACAGGCACATGAACCTGACACGACCAACAGTACTTCGACGAGC	2522						
Qy	601	AGTAACGGAAACGCAAAACGAATGACATCGGATGTTTCGGCGCTCTTCAAGAAATACAATTC	660						
Db	2523	AACAACGGCACCCAGACCAACGACACCGGCTGCTTCGGCGCCCTTCGAGGAGTACAACAGC	2582						
Qy	661	ACAAGAACCAAAACATGTGCTTCGTCGTAATAATACCTCCACCACTGCCCCACAGCCGTCGG	720						

Db	2583	ACCAGNACAGACCTGCGCCCCCAGCAAGATCCCGACGCCCTTGCACCAGCGCAGGCC	264
Qy	721	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCGGACCCCAAGC	780
Db	2643	GAGATCAAGCCACCGACACCCCCACCGAGCGCCACCACTGAAACCAACCGACCCCAAC	2702
Qy	781	AGTGATGATGAGACCTCGCAACATCGGCTCAGGGTCCGAGNACGAGAACCCACACA	840
Db	2703	AACGACGACGAGGACCTGTATCACCGCGGCGAGCGCGGAGCAGGAGCCCTACACC	2762
Qy	841	ACTTCTGATCGGCTCAACAGCAAGGCGTTTCATCAACAATGCCACCACATCCCTCACCA	900
Db	2763	ACCAGCGAGCGCTGACCAAGCAGGCGCTGAGCAGCACCATGCTCTCTACCCCTAGCCCT	2822
Qy	901	CAACCAAGCAGCGCACGACGAGGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT	960
Db	2823	CAGCCGAGCACCCCTCAGCAGGAGGGGCAACAACAACCGACACGACGAGGCGCGTGACC	2882
Qy	961	GAACCTAGACAAAATAACACAACCTGCAACAACCGTTCATGCCCCCTCATAACTACCA	1020
Db	2883	GAGCCCAACGAGACCAACACACCGCCCGCCAGCATGCTCTCTCAACAACCCAGCCGC	2942
Qy	1021	ATCTCTACTAACCAACACCTCTCAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAC	1080
Db	2943	ATCAGCACCAACAACACCCAGCAAGAACAATTCAGCACCCCTGAGCGTGAGCCTGCAGAAC	3002
Qy	1081	ACACCATGACAAACACACAGAGACACATCACTGAAAAATGAGCAAAACGAGTGCCTCCG	1140
Db	3003	ACCACCAACTACGACACCCAGAGCACCGCCACCGAGAACGAGCAGACGAGCGCCCTAGC	3062
Qy	1141	ATAACAAACCTGCTCCCAACGGGAAATCCCAACACAGCAAGAGAGCAGCAGCAAAAAA	1200
Db	3063	AGACCAACCTGCTCTCCACCGGCACTTGACACACCGCAAGAGACCAACAACACCAAG	3122
Qy	1201	GGCCCCGCCACAACGGCACCAAAACAGCAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCA	1260
Db	3123	GGCCCCACCAACACCGGCCCTTAACATGACCAACGGCCACCTGACCAAGCCCAAGCCCA	3182
Qy	1261	CCAGCTCGACTGCACACATCTTGATATTTTTCAGAGAAAGCG	1304
Db	3183	CCCAACCCCAACCCCAACCACTGGTGACTTTCAGGAAGAAGAG	3226

RESULT 13

US-10-066-506A-11

; Sequence 11, Application US/10066506A

; Publication No. US20030108560A1

; GENERAL INFORMATION:

; APPLICANT: Grogan, Case C.

; APPLICANT: Hevey, Michael C.

; APPLICANT: Schmaljohn, Alan, L.

; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein

; FILE OF INVENTION: 003/243/SAP

; CURRENT APPLICATION NUMBER: US/10/066,506A

; PRIORITY FILING DATE: 2002-01-31

; PRIOR FILING DATE: 60/267,522

; PRIOR FILING DATE: 2001-01-31

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Apple Macintosh Microsoft Word 6.0

; SEQ ID NO 11

; LENGTH: 2039

; TYPE: DNA

; ORGANISM: Ebola virus Zaire strain

; FEATURE:

; OTHER INFORMATION: chimeric molecule between Ebola virus Glycoprotein 1 and

; OTHER INFORMATION: Glycoprotein 2

US-10-066-506A-11

Query Match 29.3%; Score 540; DB 15; Length 2039;
Best Local Similarity 99.1%; Pred. No. 3.5e-157;
Matches 543; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QV 1294 AGAAGAAAGCGATCGGCAATGTGTCATGTCTCAACCCAAATGCAACCCCAATTTTCATTAC 1353

```
Db 1492 AGAACTCGACGATCGCAATTTGCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1551
Qy 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1413
Db 1552 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1611
Qy 1414 GCAGCCGAGGGAATTTACATAGAGGGGCTAAATGCACAATCAAGATGGTTAATCTGTGGG 1473
Db 1612 GCAGCCGAGGGAATTTACATAGAGGGGCTAAATGCACAATCAAGATGGTTAATCTGTGGG 1671
Qy 1474 TTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
Db 1672 TTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1731
Qy 1534 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTCTTCTGCTGAGCGATGG 1593
Db 1732 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTCTTCTGCTGAGCGATGG 1791
Qy 1594 GCGGCACATGCCACATTTCTGGACCGGACTGCTGTATCGAACCCACATGATTGGACCAAG 1653
Db 1792 GCGGCACATGCCACATTTCTGGACCGGACTGCTGTATCGAACCCACATGATTGGACCAAG 1851
Qy 1654 AACATAACAGACAAATTCATCAGATTATTCATGATTTTGTGATAAAACCCCTTCGGAC 1713
Db 1852 AACATAACAGACAAATTCATCAGATTATTCATGATTTTGTGATAAAACCCCTTCGGAC 1911
Qy 1714 CAGGGGCAATGACAAATTTGGTGACAGATGGAGACATGATACCGGCGATTTGGA 1773
Db 1912 CAGGGGCAATGACAAATTTGGTGACAGATGGAGACATGATACCGGCGATTTGGA 1971
Qy 1774 GTTACAGCGGTTATTAATTCAGTTATTCGCTTTTATCTGTATATGCAAAATTTGCTTTTAG 1833
Db 1972 GTTACAGCGGTTATTAATTCAGTTATTCGCTTTTATCTGTATATGCAAAATTTGCTTTTAG 2031
Qy 1834 TTGAATTC 1841
Db 2032 TTGAATTC 2039
```

```
RESULT 14
US-10-860-878-4
; Sequence 4, Application US/10860878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
; FILE REFERENCE: 2115a-001474POA
; CURRENT APPLICATION NUMBER: US/10/860,878
; PRIOR FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/09/600,766
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: USSN 60/072033
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 7272
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; OTHER INFORMATION: containing DNA for soluble GP of Ebola Virus,
; OTHER INFORMATION: Zaire strain
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (886)..(1129)
; OTHER INFORMATION: CMV IE 5' UT
; FEATURE:
; NAME/KEY: intron
```

```
; LOCATION: (1130)..(1840)
; OTHER INFORMATION: CMV IE INT
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: (248)..(885)
; OTHER INFORMATION: CMV enhancer
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1870)..(4288)
; OTHER INFORMATION: SGP (Z)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4289)..(4841)
; OTHER INFORMATION: TSGH
; FEATURE:
; NAME/KEY: gene
; LOCATION: (6337)..(6959)
; OTHER INFORMATION: Kan r
US-10-860-878-4

Query Match      28.7%; Score 527.6; DB 22; Length 7272;
Best Local Similarity 98.3%; Pred. No. 5.7e-153;
Matches 533; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1294 AGAAGAAAGCGATCGGCAATTTGTCATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1353
Db 3560 AGAATCTGAAGAGGAAATTTGTCATGCTCAACCCAAATGCAACCTTAATTTACATTAC 3619
Qy 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1413
Db 3620 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 3679
Qy 1414 GCAGCCGAGGGAATTTACATAGAGGGGCTTAATGCACAATCAAGATGGTTAATCTGTGGG 1473
Db 3680 GCAGCCGAGGGAATTTACATAGAGGGGCTTAATGCACAATCAAGATGGTTAATCTGTGGG 3739
Qy 1474 TTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
Db 3740 TTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 3799
Qy 1534 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTCTTCTGCTGAGCGATGG 1593
Db 3800 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTCTTCTGCTGAGCGATGG 3859
Qy 1594 GCGGCACATGCCACATTTCTGGACCGGACTGCTGTATCGAACCCACATGATTGGACCAAG 1653
Db 3860 GCGGCACATGCCACATTTCTGGACCGGACTGCTGTATCGAACCCACATGATTGGACCAAG 3919
Qy 1654 AACATAACAGACAAATTTGATCAGATTATTCATGATTTTGTGATAAAACCCCTTCGGAC 1713
Db 3920 AACATAACAGACAAATTTGATCAGATTATTCATGATTTTGTGATAAAACCCCTTCGGAC 3979
Qy 1714 CAGGGGCAATGACAAATTTGGTGACAGATGGAGACATGATACCGGCGATTTGGA 1773
Db 3980 CAGGGGCAATGACAAATTTGGTGACAGATGGAGACATGATACCGGCGATTTGGA 4039
Qy 1774 GTTACAGCGGTTATTAATTTGCAAGTTATTCGCTTTTATCTGTATATGCAAAATTTGCTTTTAG 1833
Db 4040 GTTACAGCGGTTATTAATTTGCAAGTTATTCGCTTTTATCTGTATATGCAAAATTTGCTTTTAG 4099
Qy 1834 TT 1835
Db 4100 TT 4101

RESULT 15
US-10-860-878-3
; Sequence 3, Application US/10860878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
```

;; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
;; TITLE OF INVENTION: By Pseudotyping With Viral Glycoprotein
;; FILE REFERENCE: 21158-001474FOA
;; CURRENT APPLICATION NUMBER: US/10/860,878
;; CURRENT FILING DATE: 2004-06-03
;; PRIOR APPLICATION NUMBER: US/09/600,766
;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: USSN 60/072033
;; PRIOR FILING DATE: 1998-01-21
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 7285
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Plasmid
;; OTHER INFORMATION: containing DNA of GP of Ebola Virus, Zaire strain
;; FEATURE:
;; NAME/KEY: 5'UTR
;; LOCATION: (686)..(1129)
;; OTHER INFORMATION: CMV IE 5' UT
;; FEATURE:
;; NAME/KEY: intron
;; LOCATION: (1130)..(1840)
;; OTHER INFORMATION: CMV IE INT
;; FEATURE:
;; NAME/KEY: enhancer
;; LOCATION: (248)..(885)
;; OTHER INFORMATION: CMV enhancer
;; FEATURE:
;; NAME/KEY: gene
;; LOCATION: (1870)..(4301)
;; OTHER INFORMATION: GP(Z)
;; FEATURE:
;; NAME/KEY: gene
;; LOCATION: (4302)..(4854)
;; OTHER INFORMATION: TGGH
;; FEATURE:
;; NAME/KEY: gene
;; LOCATION: (6350)..(6972)
;; OTHER INFORMATION: Kan r
US-10-860-878-3

Query Match 28.7%; Score 527.6; DB 22; Length 7285;
Best Local Similarity 98.3%; Pred. No. 5.7e-153;
Matches 533; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1294 AGAAGAACCGATCGGCAATTGTCATGCTCAACCCAAATGCCAACCCCTAAATTACATTAC 1353
DB 3563 AGAACTCGAAGAGAGCAATTGTCATGCTCAACCCAAATGCCAACCCCTAAATTACATTAC 3622
QY 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTGGATACCATATTTCCGGGCA 1413
DB 3623 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTGGATACCATATTTCCGGGCA 3682
QY 1414 GCAGCGAGGGAATTTACATAGAGGGCTAATGCAATCAAGATGGTTTAAATCTGTGG 1473
DB 3683 GCAGCGAGGGAATTTACATAGAGGGCTAATGCAATCAAGATGGTTTAAATCTGTGG 3742
QY 1474 TTGAGACAGCTGGCCACGAGACGACTCAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
DB 3743 TTGAGACAGCTGGCCACGAGACGACTCAGCTCTTCAACTGTTCTGAGAGCCCAACT 3802
QY 1534 GAGCTACGCACTTTTCAATCTCAACCGTAAGGCAATTTGTTCTGTCGAGCGATGG 1593
DB 3803 GAGCTACGCACTTTTCAATCTCAACCGTAAGGCAATTTGTTCTGTCGAGCGATGG 3862
QY 1594 GCGGCACATGCCACATTTCTGGGACCGGACTGCTGTATCGAACCCATGATTGGACCAAG 1653
DB 3863 GCGGCACATGCCACATTTCTGGGACCGGACTGCTGTATCGAACCCATGATTGGACCAAG 3922
QY 1654 AACATAACAGCAAAATTCATGATTATTATGATGTTTGTGATAAAACCCCTCCGGAC 1713

DB 3923 AACATAACAGCAAAATTCATGATTATTATCATGATTTTGTGATAAAACCCCTCCGGAC 3982
QY 1714 CAGGGGACCAATGACAATTTGGTGGACAGGATGGAGACAATGGATACCGGCAAGTATTGGA 1773
DB 3983 CAGGGGACCAATGACAATTTGGTGGACAGGATGGAGACAATGGATACCGGCAAGTATTGGA 4042
QY 1774 GTTACAGCGCTTATAAATTCGAGTTATCGCTTTATTCTGTATATGCAAAATTTGCTTTTAG 1833
DB 4043 GTTACAGCGCTTATAAATTCGAGTTATCGCTTTATTCTGTATATGCAAAATTTGCTTTTAG 4102
QY 1834 TT 1835
DB 4103 TT 4104

Search completed: August 1, 2005, 20:40:59
Job time : 1161.01 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 8793.18 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacattgttcct.....tcactaaacattgatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2046	100.0	2046	6	AX717717	AX717717 Sequence
2	1841.2	90.0	11460	6	AR279677	AR279677 Sequence
3	1841.2	90.0	19104	14	MVREPCYC	Z12132 Marburg vir
4	1841.2	90.0	19112	14	AY430366	AY430366 Lake Vict
5	1841.2	90.0	19113	14	AY430365	AY430365 Lake Vict
6	1813.6	88.6	2051	6	AX717721	AX717721 Sequence
7	1695.6	82.9	2046	14	AF005735	AF005735 Marburg v
8	1692.4	82.7	2948	14	MAVSPAA	X68493 Marburg vir
9	1692.4	82.7	19112	14	MVIRPR	Z29337 Marburg vir
10	1653.6	80.8	19151	14	AY358025	AY358025 Marburg v
11	1652.4	80.8	2046	14	AF005733	AF005733 Marburg v
12	1518	74.2	2046	6	AX717725	AX717725 Sequence
13	1513.2	74.0	2046	14	AF005734	AF005734 Marburg v
14	1319.6	64.5	2046	6	AX717719	AX717719 Sequence
15	1308	63.9	1841	6	AX717715	AX717715 Sequence
16	546.6	26.7	2252	6	AX717713	AX717713 Sequence
17	148.6	7.3	2261	14	AY344234	AY344234 Sudan ebo
18	148.6	7.3	2363	14	AY316199	AY316199 Sudan ebo
19	148.6	7.3	18875	14	AY729654	AY729654 Sudan ebo

20	145	7.1	2362	14	EVU23069	U23069 Sudan Ebola
21	145	7.1	2362	14	EVU28134	U28134 Sudan Ebola
22	138.8	6.8	2408	14	EVU28006	U28006 Cote d'Ivoire
23	134.4	6.6	2408	14	EVU81161	U81161 Zaire Ebola
24	132.8	6.5	18959	14	AY142960	AY142960 Zaire Ebo
25	132.8	6.5	18960	14	AF499101	AF499101 Zaire Ebo
26	131.2	6.4	2039	6	AX717723	AX717723 Sequence
27	131.2	6.4	2173	14	AY058898	AY058898 Zaire Ebo
28	131.2	6.4	2298	6	AR404895	AR404895 Sequence
29	131.2	6.4	2298	6	AX092102	AX092102 Sequence
30	131.2	6.4	2406	14	EVU31033	U31033 Zaire Ebola
31	131.2	6.4	2408	14	EVU23187	U23187 Zaire Ebola
32	131.2	6.4	2408	14	EVU28077	U28077 Zaire Ebola
33	131.2	6.4	11742	14	EBORNA	IL1365 Zaire Ebola
34	131.2	6.4	18959	14	AF086833	AF086833 Zaire Ebo
35	131.2	6.4	18959	14	AF272001	AF272001 Zaire Ebo
36	131.2	6.4	18961	14	AY354458	AY354458 Zaire ebo
37	129.6	6.3	2173	14	AY526098	AY526098 Zaire ebo
38	129.6	6.3	2173	14	AY526100	AY526100 Zaire ebo
39	129.6	6.3	2173	14	AY526101	AY526101 Zaire ebo
40	129.6	6.3	2446	14	EVU77384	U77384 Zaire Ebola
41	128	6.3	1041	14	AY526103	AY526103 Zaire ebo
42	128	6.3	2173	14	AY526099	AY526099 Zaire ebo
43	128	6.3	2173	14	AY526102	AY526102 Zaire ebo
44	128	6.3	2173	14	AY526104	AY526104 Zaire ebo
45	128	6.3	2173	14	AY526105	AY526105 Zaire ebo

ALIGNMENTS

RESULT 1
AX717717
LOCUS AX717717 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Musoke Glycoprotein 1 and Marburg virus strain Raven
Glycoprotein 2"

ORIGIN	Query Match	100.0%;	Score 2046;	DB 6;	Length 2046;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2046;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGACACAAAAATCTCCCC	60		
Db	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGACACAAAAATCTCCCC	60		
Qy	61	ATTTTAGAGTAGCTAGTAAATCAACCCCAAAATGGATTCCGTTATGTCCTCGGAAT	120		
Db	61	ATTTTAGAGTAGCTAGTAAATCAACCCCAAAATGGATTCCGTTATGTCCTCGGAAT	120		
Qy	121	CTCCAGAACAGACAGACGTCCTATCTGATGGGATTCACCTGAGTGGGCAAAAAGTTGCT	180		
Db	121	CTCCAGAACAGACAGACGTCCTATCTGATGGGATTCACCTGAGTGGGCAAAAAGTTGCT	180		
Qy	181	GATTCCCTTTGGAGGAGATCCACCGATGGGCTTTCAGGACAGGTGTACTCTCCCAAGNAT	240		

Db 181 |||||GATTCCTCCCTTTGGAGGCATCCAAGGCGATGGGCTTTTCAGACAGAGTGTTACCTCCCAAGAA 240
Qy 241 GTTGTAGTACACAGAGGGGAGGAGCCAAACATGCTCAATATATAGTGTACGGATCCC 300
Db 241 GTTGTAGTACACAGAGGGGAGGAGCCAAACATGCTCAATATATAGTGTACGGATCCC 300
Qy 301 TCTGGAATCCTTGTCTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAATCCTTGTCTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
Qy 361 ACTATCCATCATATTCAGGTCGCTATGATGCTCCCAACAACTGCTACCGAGGCAAGCTTC 420
Db 361 ACTATCCATCATATTCAGGTCGCTATGATGCTCCCAACAACTGCTACCGAGGCAAGCTTC 420
Qy 421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAACTGCTACCGAGGCAAGCTTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAACTGCTACCGAGGCAAGCTTC 480
Qy 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
Qy 541 CGGCAAGCAAGGATACCGTCAATATGATCTGACTTCTAATAAATATTTGGACAAGT 600
Db 541 CGGCAAGCAAGGATACCGTCAATATGATCTGACTTCTAATAAATATTTGGACAAGT 600
Qy 601 AGTAACGGAACCAACGAATGACACTGGATGTTTCGGGCTCTTCAAGATACAAATCT 660
Db 601 AGTAACGGAACCAACGAATGACACTGGATGTTTCGGGCTCTTCAAGATACAAATCT 660
Qy 661 ACAAGAACCAACATGTGCTCCGTCCTCAAAATACCTCCACACATGCTCCACAGCCGTCG 720
Db 661 ACAAGAACCAACATGTGCTCCGTCCTCAAAATACCTCCACACATGCTCCACAGCCGTCG 720
Qy 721 GAGATCAAACTCAAGCAACCCAACTGATGCTCAACAACTCAATACACAGGACCCCAAGC 780
Db 721 GAGATCAAACTCAAGCAACCCAACTGATGCTCAACAACTCAATACACAGGACCCCAAGC 780
Qy 781 AGTGATGAGGACCTCGCAACATCCGGCTCAGGTCGGAGACGAGACCCCAACA 840
Db 781 AGTGATGAGGACCTCGCAACATCCGGCTCAGGTCGGAGACGAGACCCCAACA 840
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGCTTTTCATCAACAACTGCCACCACTCCCTCACCA 900
Db 841 ACTTCTGATGCGGTCAACAGCAAGGCTTTTCATCAACAACTGCCACCACTCCCTCACCA 900
Qy 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACAAACCAATTCGCCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACAAACCAATTCGCCAAGATGCTGTGACT 960
Qy 961 GAACTAGACAAAATAACACAACTGCAACCGTCCATGCCCTCATTAACACTACCACA 1020
Db 961 GAACTAGACAAAATAACACAACTGCAACCGTCCATGCCCTCATTAACACTACCACA 1020
Qy 1021 ATCTCTACTAACAACACCTCCAAACACACACTTTCAGCACTCTCTGCACTTACAAAC 1080
Db 1021 ATCTCTACTAACAACACCTCCAAACACACACTTTCAGCACTCTCTGCACTTACAAAC 1080
Qy 1081 ACCACAATGACACACAGAGCAATCACTGAAAATGAGCAAAACGATGCCCTCCG 1140
Db 1081 ACCACAATGACACACAGAGCAATCACTGAAAATGAGCAAAACGATGCCCTCCG 1140
Qy 1141 ATAAACAACCTGCTCCCAAGGGAATCCCAACAGCAAGGACCCAGCAGCAAAA 1200
Db 1141 ATAAACAACCTGCTCCCAAGGGAATCCCAACAGCAAGGACCCAGCAGCAAAA 1200
Qy 1201 GGGCCCGCCACACAGGACCAACAGCAATGAGCAATTTCCAGCTCTCCCGCCACC 1260
Db 1201 GGGCCCGCCACACAGGACCAACAGCAATGAGCAATTTCCAGCTCTCCCGCCACC 1260
Qy 1261 CCCAGCTGACGTGCAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320

Db 1261 CCCAGCTGACGTGCAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
Qy 1321 GAAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Db 1321 GAAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Qy 1381 ATCCAAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db 1381 ATCCAAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Qy 1441 GAAACAACACTCCCGGAATATCAGTTTAACTTTCTTCTTATTTTCTGATAAATAATGGA 1500
Db 1441 GAAACAACACTCCCGGAATATCAGTTTAACTTTCTTCTTATTTTCTGATAAATAATGGA 1500
Qy 1501 GATACTGCTACTCTGGGGAACGAGAAATGATTTGATGACAGAGTTGAGGATTTGGAGT 1560
Db 1501 GATACTGCTACTCTGGGGAACGAGAAATGATTTGATGACAGAGTTGAGGATTTGGAGT 1560
Qy 1561 GTGAGAGGACGATTTGGCGGAGGCTTACGTAGTGATGATGATGATGATGATGATGATGATG 1620
Db 1561 GTGAGAGGACGATTTGGCGGAGGCTTACGTAGTGATGATGATGATGATGATGATGATGATG 1620
Qy 1621 GAAAGCTCTATATCTGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGT 1680
Db 1621 GAAAGCTCTATATCTGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGT 1680
Qy 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACCGAGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACCGAGAA 1740
Qy 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGATGATGATGATGATGATGATGATGATG 1800
Db 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGATGATGATGATGATGATGATGATGATG 1800
Qy 1801 ACATGCAAGGTCAGGACCTGATTTGTCATAGGATAGAGATCTATCTAATAATATC 1860
Db 1801 ACATGCAAGGTCAGGACCTGATTTGTCATAGGATAGAGATCTATCTAATAATATC 1860
Qy 1861 TCAGAACAAATCGACAAATCAGAAAGGATCAACAAAGGAGGAACTGGCTGGGCTCTA 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATCAACAAAGGAGGAACTGGCTGGGCTCTA 1920
Qy 1921 GGTGGCAATCGTGACATCTGACTGGGGTGTCTCACCAATTTGGGCACTCTGCTACTA 1980
Db 1921 GGTGGCAATCGTGACATCTGACTGGGGTGTCTCACCAATTTGGGCACTCTGCTACTA 1980
Qy 1981 TTATCTATAGTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db 1981 TTATCTATAGTGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Qy 2041 GGATGA 2046
Db 2041 GGATGA 2046

RESULT 2
AR279677
LOCUS AR279677 11460 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6517842.
ACCESSION AR279677
VERSION AR279677.1 GI:29714600
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 11460)
AUTHORS Hevey,M.C., Negley,D.L., Pushko,P., Smith,J.F. and Schmaljohn,A.L.
TITLE Marburg virus vaccines
JOURNAL Patent: US 6517842-A 1 11-FEB-2003;
FEATURES Location/Qualifiers
 1..11460
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 90.0%; Score 1841.2; DB 6; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTCTATAGTCTTATCTTAATTTCAAGGACAAAAATCTCCC 60
DB |||||
QY 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCTCGTATGCTCCGGAAT 120
DB |||||
QY 6000 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCTCGTATGCTCCGGAAT 6059
DB |||||
QY 121 CTCGAGAGACAGAGAGAGCTCATCTGATGGATTTCACATGAGTGGGCAAAAGTTGCT 180
DB |||||
QY 6060 CTCGAGAGACAGAGAGAGCTCATCTGATGGGATTTCACATGAGTGGGCAAAAGTTGCT 6119
DB |||||
QY 181 GATTCCTCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
DB |||||
QY 6120 GATTCCTCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 6179
DB |||||
QY 241 GTTAGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGTAAACGGATCCC 300
DB |||||
QY 6180 GTTAGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGTAAACGGATCCC 6239
DB |||||
QY 301 TCTGAAAATCTTCTGTTAGATCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
DB |||||
QY 6240 TCTGAAAATCTTCTGTTAGATCTCTACCAACATCCGTGACTATCCGAAATGCAAA 6299
DB |||||
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB |||||
QY 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 6359
DB |||||
QY 421 GGAGCATTTTCTGTATGATCGCATTCCTCCACAAATGATGACCGAGGCAAAAGTCTTC 480
DB |||||
QY 6360 GGAGCATTTTCTGTATGATCGCATTCCTCCACAAATGATGACCGAGGCAAAAGTCTTC 6419
DB |||||
QY 481 ACTGAAGGGAACATAGCAGCTATGATGTCATAAGACAGATGACCAAAATGATTTTCTCG 540
DB |||||
QY 6420 ACTGAAGGGAACATAGCAGCTATGATGTCATAAGACAGATGACCAAAATGATTTTCTCG 6479
DB |||||
QY 541 CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTACTAATAATATTTGGCAAGT 600
DB |||||
QY 6480 CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTACTAATAATATTTGGCAAGT 6539
DB |||||
QY 601 AGTAAACGGAACCAACGAATGACATGATGTTTCGGCGCTTTTCAAGAAATACAAATTTCT 660
DB |||||
QY 6540 AGTAAACGGAACCAACGAATGACATGATGTTTCGGCGCTTTTCAAGAAATACAAATTTCT 6599
DB |||||
QY 661 ACAAGAAACCAACATGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 720
DB |||||
QY 6600 ACAAGAAACCAACATGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 6659
DB |||||
QY 721 GAGATCAAACTCAAGACACCCCACTGATGACCAAACTCAATACACGACCCCAAGC 780
DB |||||
QY 6660 GAGATCAAACTCAAGACACCCCACTGATGACCAAACTCAATACACGACCCCAAGC 6719
DB |||||
QY 781 AGTGTATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGACGAGAACCCCAACACA 840
DB |||||
QY 6720 AGTGTATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGACGAGAACCCCAACACA 6779
DB |||||
QY 841 ACTTCTGATGCGGTCAACGAACAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 900
DB |||||
QY 6780 ACTTCTGATGCGGTCAACGAACAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 6839
DB |||||
QY 901 CAACCAAGCAGCGCACAGCAAGGAGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 960
DB |||||
QY 6840 CAACCAAGCAGCGCACAGCAAGGAGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 6899
DB |||||
QY 961 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCGCCCTCATTAACACTACCACA 1020
DB |||||
QY 6900 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCGCCCTCATTAACACTACCACA 6959
DB |||||

QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGACCACTTACAAAC 1080
DB |||||
QY 6960 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGACCACTTACAAAC 7019
DB |||||
QY 1081 ACCACCAATGACACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTGCCTCCG 1140
DB |||||
QY 7020 ACCACCAATGACACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTGCCTCCG 7079
DB |||||
QY 1141 ATAAACACCTTCCTCCAAACCGGAAATCCCAACACAGCAAAAGAGACACAGAGCAAAAA 1200
DB |||||
QY 7080 ATAAACACCTTCCTCCAAACCGGAAATCCCAACACAGCAAAAGAGACACAGAGCAAAAA 7139
DB |||||
QY 1201 GGCCCGCCCAACCGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCGCCAC 1260
DB |||||
QY 7140 GGCCCGCCCAACCGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCGCCAC 7199
DB |||||
QY 1261 CCCAGCTCGATGACACAACTCTTTGTATATTTTCAGAAAGAGCGATCGATTTCTGGAAA 1320
DB |||||
QY 7200 CCCAGCTCGATGACACAACTCTTTGTATATTTTCAGAAAGAGCGAGTATCTCTGGAGG 7259
DB |||||
QY 1321 GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAAATACTGAAATTTGATTTGATCCA 1380
DB |||||
QY 7260 GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAAATACTGAAATTTGATTTGACCCA 7319
DB |||||
QY 1381 ATCCCAACACAGAAAACATCTTTGATGATCTCCAGCTTTAATACTTCAACTAATGAG 1440
DB |||||
QY 7320 GTTCCAAATACAAAACAACTTTGATGATCTCTAGTCTCTGGTGCCTCGGCTGAGGAA 7379
DB |||||
QY 1441 GAAACAAACACTCCCGGAAATATCACTTTTAACTTTCTCTATTTTCTGATAAAATGGA 1500
DB |||||
QY 7380 GATCAACATGCTCCCGCAATATAGTTTAACTTTATCTTATTTCTTAATAATAATGAG 7439
DB |||||
QY 1501 GATCTGCTACTCTCTGGGAAAAACGAGAAATGATGATGATGAGAGTTGAGGATTTGGAGT 1560
DB |||||
QY 7440 AACCTGCTACTCTCTGGAGAAATGAGAAATGATGATGATGAGAGTTAAGAAATTTGGAGC 7499
DB |||||
QY 1561 GTGAGGAGGACGATTTGCGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1620
DB |||||
QY 7500 GTTCAGGAGGATGACCTGCGCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGAAT 7559
DB |||||
QY 1621 GAAAGACTCTACTGCGCGTTTTTAATCAAAATCAAGAAATTTAGTTTGTAGTTTGGG 1680
DB |||||
QY 7560 GAAAGACTCTACTGCTGTTTTTAATAAAATCAAAACAAATTTGCTGCAAGTTGAGG 7619
DB |||||
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTAAGGGTCAACACGAGGAA 1740
DB |||||
QY 7620 CGCTAGCCAATCAAACTGCCAAATCTTTGGAATCTTTATGAGAGTCAACACTGAGGAA 7679
DB |||||
QY 1741 AGGACATTTTCTTAATCAATAGGACGCAATTTGATCTTTTGTCTTACGAGGTGGGGCGGA 1800
DB |||||
QY 7680 AGAACATTTCTCTTAATCAATAGACATGCTATTTGACTTTTACTCACAAGATGGGGAGGA 7739
DB |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
DB |||||
QY 7740 ACATGCAAGGTGCTTGGACCTGATTTGTCATCGGATAGAGACTTTGTCCAAAATATT 7799
DB |||||
QY 1861 TCAGAACAAATCGACAAAATCAGAAAGGATGAAACAAAAAGGAGGAACTGGCTGGGGTCTA 1920
DB |||||
QY 7800 TCAGAGCAATTTGACCAATTTAAAGAGGACGAAACAAAAAGAGGGGACTGGTTGGGGTCTG 7859
DB |||||
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGATCTCTGTACTA 1980
DB |||||
QY 7860 GGTGTAATGGTGACATCTGACTGGGGTGTCTTACTAATTTGGGATTTTGTACTA 7919
DB |||||
QY 1981 TTATCTATAGTGTCTGATGCTCTGCTGATCTCTGCTGATCTTCACTAAATACATT 2040
DB |||||
QY 7920 TTATCTATAGTGTCTGATGCTCTGCTGATCTCTGCTGATCTTACTAAATATATC 7979
DB |||||
QY 2041 GGATGA 2046
DB |||||
QY 7980 GGATGA 7985
DB |||||

gene	APENPLNGPNNRQVVLAYANPTLSAV"	181	GATTCCCTTTGGAGGATCCAAAGCATGGCTTTGAGGACAGGTGTACCTCCCAAGAAT	240
mRNA	5822..8666	Db		6120
	/gene="gp"			
	5822..8666			
	/gene="gp"			
	/standard_name="gene four"			
	/citation=[3]			
	/evidence=experimental			
CDS	5940..7985			
	/gene="gp"			
	/standard_name="gp"			
	/citation=[3]			
	/codon_start=1			
	/evidence=experimental			
	/product="glycoprotein"			
	/protein_id="CAA78117.1"			
	/db_xref="GI:60627"			
	/db_xref="GOA:P35253"			
	/db_xref="UniProt/Swiss-Prot:P35253"			
	/translation="MKTTCFLSLILIGTKNLPFILETASNNQPNQVDSVCSGTLOKT EDVHLMGFTLSGQKVADSPLEASKRWAFRTGPPKNVEYTEGEBAKTCYINISVDPDSG KSLIDPPTNIRDYPKCKTIIHQONPHQAGIALHLWGAFLYDRISTATYMRGVF TEGNTAAMVNTKTHMIFSRQOGYRHMNLTSTNKWTSNGTQNDTGTGFGALOEY NSTKNOTCAPSKIPEPLTARPEIKLTSTPTDACKLNTTDPSSDEDLATSGSGGER EPHTSDAVTKQSLSTWPTSPQSPPTQOGGNTHSQDAVTELDKNWTAQPSMP PHNTTISTNNRSHNFSLAPLQNTNDNTQSTITENEQTSAPSITLLEPPTGNPT AKSTSSKGPATAPNTNEHTSPPTPSAQHLVYFRKRSILWREGDMFPFLDG LINAPIDPVPNTKTIPESSSSGASAEQHASPNISLTLSPFNINENTAYSGEN ENDCDALRIVSWOEDDLAAGLSWIPFPFGIEGLYTAVLTKNQNNLVCRURRANOT AKSLELLRVTEERTFSLNRHAIDPFLTRWGTCKVLGPDCCIGIEDLSKNTSEI DQIKDEQKBTGHWLGSKWNTSDWGLTNLIGILLLSIAVLIALSCICRIPTKYIG" 8762..10009			
gene	/gene="vp30"			
mRNA	8762..10009			
	/gene="vp30"			
	/standard_name="gene five"			
	/citation=[3]			
	/evidence=experimental			
	8864..9697			
	/gene="vp30"			
	/standard_name="vp30"			
	/citation=[3]			
	/codon_start=1			
	/evidence=experimental			
	/product="vp30"			
	/protein_id="CAA78118.1"			
	/db_xref="GI:541782"			
	/db_xref="GOA:P35258"			
	/db_xref="UniProt/Swiss-Prot:P35258"			
	/translation="MQQPRGRSRTRNHQVPTIYHETQLPSKPHYNTVHPRAKMSST RSSAESPTNHI PRARPPSTFNLSKPPPPKDMCRNMKIGLPCADPTCNDRDIDLNT NRELLLMARKMLPNTDKTFRMPQDCGSPSLSKGLSKDKQBTQDVLTLNGLHILSY LHRSFIGNMRHLRAALSLTCAGIRKTRNSLINTWTELMHNLPODQNGVIKQYIT GHLDKGQGFENALMQGWDKRSISLIFVQAALYVMNNIFCESSISVQASYESFYSSKS R" 9990..11278			
gene	/gene="vp24"			
	Query Match 90.0%; Score 1841.2; DB 14; Length 19104;			
	Best Local Similarity 93.7%; Pred. No. 0; Mismatches 128; Indels 0; Gaps 0;			
	Matches 1918; Conservative 0;			
QY	1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTTAAATTCAGGGACAAAAATCTCCCC 60			
Db	5940 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTTAAATTCAGGGACAAAAATCTCCCC 5999			
QY	61 ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120			
Db	6000 ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059			
QY	121 CTCGAGACAGACAGACCTGTCATCTGTGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180			
Db	6060 CTCGAGACAGACAGACCTGTCATCTGTGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119			

QY	1261	CCAGCTCGACTGCACAAACATCTTGTAATATTTTCAGAGAAAGCGATTCGATTTCTTGGAAA	1320
Db	7200	CCAGCTCGACTGCACAAACATCTTGTAATATTTTCAGAGAAAGCGAAGTATCTCTCGGAGG	7259
QY	1321	GAAGTGATATATCCCGTTTTAGATGGTTAAATATACTGAATGATTTGATTCATCCA	1380
Db	7260	GAAGCGACATGTTCCCTTTCTCGATGGTTAAATATACTGAATGATTTGATTCATCCA	7319
QY	1381	ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAAATACCTTCAACTAATGAG	1440
Db	7320	GTTCCAATACAAACAAATCTTTGATGAATCTCTAGTCTGGTCCCTCGCTGAGGAA	7379
QY	1441	GAACAACACATCCCGCGAATACAGTTAACTTCTTATTTCTGATATAAAATGGA	1500
Db	7380	GATCAACATGCTCCCGCAATATAGTTAACTTATCTTATTTCTTAAATATAATGAG	7439
QY	1501	GATAGTCTACTCTGGGGAACGAGATGATTCGATGACAGTTGAGGATTTGGAGT	1560
Db	7440	AACACTGCTCTCTGGAGAAAATGAGATGATTTGATGCGAGGTTAAGAAATTTGGAGC	7499
QY	1561	GTGCAGGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC	1620
Db	7500	GTTCAGGAGATGACTGCGCGAGGCTCAGTTGGATACCGTTTTTGGCCCTCGAAT	7559
QY	1621	GAAGACTCTATATGCGCGGTTAATCAAAAATCAGAACAAATTTAGTTGTAGTTGAGG	1680
Db	7560	GAAGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAAATTTGCTCTGCAGGTTGAGG	7619
QY	1681	CGCTTAGCTAATCAAACTGTAATCCTTGGAGCTTGTGTAAGGTCACACCGAGGAA	1740
Db	7620	CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA	7679
QY	1741	AGGACATTTCTTAAATCAATAGGATGCAATTTGACTTTTCTTACGAGGTGGGGCGGA	1800
Db	7680	AGAACATTTCTTAAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGAGA	7739
QY	1801	ACATGCAAGGTGCTAGGACCTGATTGTTGCATAGGAATAGAGATCTATCTAAAAATATC	1860
Db	7740	ACATGCAAAAGTCTTGGACCTGATTGTTGCATCGGATAGAGACTTGTCCAAAAATATT	7799
QY	1861	TCAGAACAAATCGACAAATCAGAAAGATGAACAAAGGAGGAACTGGCTGGGGTCTA	1920
Db	7800	TCAGAGCAAAATTGACAAATTAATAAGGACGACAAAGAGGGGACCTGGTTGGGGTCTG	7859
QY	1921	CGTGGCAAAATGTTGGACATCTGACTGGGGTGTCTCACCAAATTTGGGCACTCTGCTACTA	1980
Db	7860	GTTGGTAAATGTTGGACATCCGACTGGGGTGTCTTACTTACTTGGGCAATTTGCTACTA	7919
QY	1981	TTATCTATAGCTGTTCTGATGCTCTGCTCTGATCTATCTGCTGATCTTCACTAAATACATT	2040
Db	7920	TTATCATAGCTGTTCTGATGCTCTGCTCTGATCTATCTGCTGATCTTCTACTAAATATATC	7979
QY	2041	GGATGA 2046	
Db	7980	GGATAA 7985	
RESULT 4			
AY430366			
LOCUS			
DEFINITION	Lake Victoria marburgvirus strain pp4 guinea pig nonlethal variant, complete genome.	19112 bp	RNA linear VRL 06-FEB-2004
ACCESSION	AY430366		
VERSION	AY430366.1	GI:40388379	
KEYWORDS			
SOURCE	Lake Victoria marburgvirus		
ORGANISM	Lake Victoria marburgvirus		
REFERENCE	1 (bases 1 to 19112)		
AUTHORS	Hevey, M., Negley, D., Geisbert, J., Jahrling, P., and Schmaljohn, A.		
TITLE	Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants		

JOURNAL	Virology 239 (1), 206-216 (1997)
MEDLINE	98087840
PUBMED	9426460
REFERENCE	2 (bases 1 to 19112)
AUTHORS	Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H., Smith, K.L., and McCready, P.M.
TITLE	Direct Submission
JOURNAL	Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
REFERENCE	3 (bases 1 to 19112)
AUTHORS	Ichou, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Lofets, L., Hevey, M., and Schmaljohn, A.
TITLE	Direct Submission
JOURNAL	Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort Detrick, MD 21702, USA
FEATURES	Location/Qualifiers
source	1..19112
	/organism="Lake Victoria marburgvirus"
	/mol_type="genomic RNA"
	/strain="pp4 guinea pig nonlethal variant"
	/isolate="Musoke Kenya, 1980"
	/db_xref="taxon:11269"
gene	104..2191
	/gene="NP"
CDS	104..2191
	/note="nucleoprotein"
	/codon_start=1
	/product="NP protein"
	/protein_id="AAR85453.1"
	/db_xref="GI:40388380"
	/translation="MDLSLLELGTPTAPHRNKKVILFDTHQVSIQNCIIDAIS GIDLDLLEGLLTLCEHYNSDKDFNTSPIAKYLDAGYEDVIEKADALPLDV IPNPHYSPLTALKTLESTQSRIGLFLSCFLPLKLVGDRAGIEKALROVTY HQEGIVTYPNHMTTGHKMVIKGLSSFLKVLJHOGVNLVTHGHDAYDSIISNV QTFRGLLVKINLEHGLYPOLSAIAGVATAGSTLAGVNVGQYQQLREAAHDAEV PFARVNLGSLNLEHGLYPOLSAIAGVATAGSTLAGVNVGQYQQLREAAHDAEV KLQRHGHQEIQAIAEDDEERKILEQHLQKTEITHSQTLAVLSQKREKLAAEIE NNIVDEQCFQSONRVSQSLNDPTPEVTQVQARPMNRPTALPPVDKLIKHESTEDS SSSSFDLNDPFPALLNEDEDTLDDSVIIGCTTSREFOGIPEPPRQSDLNNSQCKOE DESTNPIKQFLRYQELPPVQEDDESEYTTDSQSIDQPGSDNEGVDLPPPLAQE KRQPIQHPAANQDPFGSDGVNDGIDLEPIRSPSPSPAQEDTMRSAEYLSPTFN DEDQNMPPQVVTWKGRFTLYPNLDLQTNPPESLITALVVEYQNPVSAKELQADWPD MSFDERRHVAMNL"
gene	2945..3934
	/gene="VP35"
CDS	2945..3934
	/note="minor component of RNA-dependent RNA-polymerase"
	/codon_start=1
	/product="VP35 protein"
	/protein_id="AAR85454.1"
	/db_xref="GI:40388381"
	/translation="MWSSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRPKGTVG LQSPCLMSKATSDDDIWDQLIVKRLADLLIPINQISDIQSTLSVTTTRVHEIER QLCEITPVLMKGRTELEAISKGMSEMLAKYDLHVI STGRTPAFAADAYLNEHGVPPP QPAJFKDLGVAQAQCSKGTWKNATTDAAADKMSKVLSEETFSKPNLSAKDLALLF THLPGNNTPHILAQVLSKIAYKSGSGAFDAFHQIILSEGENAQALTRLSRTDAF LGVVPPVTRVKNQTVPRPCQKSLRAVPPNPTIDKGWVCVYSSEQGETRAULKI"
gene	4568..5479
	/gene="VP40"
CDS	4568..5479
	/note="matrix protein"
	/codon_start=1
	/product="VP40 protein"
	/protein_id="AAR85455.1"
	/db_xref="GI:40388382"
	/translation="MASSSNYNTYQVLYNPYPADHGANQLIPADOLSNQOQITPNVY GDNLDDQKGNVCHAPTLEAI IDISA YNERTVKGVPALPLGIMSNFEPYLAHTVAA LLTSSYITQFTHNGQKFPVRNLTGTGI PAHPLMLREGNQAFIQNMWIPNFTNQF TYNTUNLVLSVQKLPDDARPSKDKLIGNTHWPAVS IHPNLPPIVLPTVKQAYKQHK PNNGPLLAISGILHQLURVPEKTSIFRISLPADMFVSKEGMMKRGSGSPVYVYFQ

|||||
6541 AGTAACGGAAACGAAACGAATGACATGTGATGTTTCGGCGCTCTTCAAGAATACAAATCTCT 6600
QY
661 ACAAAGAAACAAACATGTCTCGTCCAAATACCTCCACCACTGCCACACAGCCCGTCCG 720
Db
6601 ACAAGAACCAACATGTCTCGTCCAAATACCTCCACCACTGCCACACAGCCCGTCCG 6660
QY
721 GAGATCAAACTCACAAGCAACCCCAACTGATGCGCAACAACTCAATCAACGAGCCCAAGC 780
Db
6661 GAGATCAAACTCACAAGCAACCCCAACTGATGCGCAACAACTCAATCAACGAGCCCAAGC 6720
QY
781 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGTCGGAGACGAGACCCCAACACA 840
Db
6721 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGTCGGAGACGAGACCCCAACACA 6780
QY
841 ACTTCTGATCGCGTCAACAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACCA 900
Db
6781 ACTTCTGATCGCGTCAACAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACCA 6840
QY
901 CAACCAAGCACGCCACAGCAAGGAGAAACAAACAACCAATTCCTCAAGATGCTGTGACT 960
Db
6841 CAACCAAGCACGCCACAGCAAGGAGAAACAAACAACCAATTCCTCAAGATGCTGTGACT 6900
QY
961 GAACCTAGACAAATAACAACTGCAACACCGTCCATGCGCCCTCATAACTACCACTACCACA 1020
Db
6901 GAACCTAGACAAATAACAACTGCAACACCGTCCATGCGCCCTCATAACTACCACTACCACA 6960
QY
1021 ATCTCTACTAAACAACACTCCAAACCAACTTCAGCACTCTCTGCACTTACCAAAAC 1080
Db
6961 ATCTCTACTAAACAACACTCCAAACCAACTTCAGCACTCTCTGCACTTACCAAAAC 7020
QY
1081 ACCACCAATGACAACACAGAGCACAATCACTGAAAATGAGCAACCAAGTGCCTCCCTCG 1140
Db
7021 ACCACCAATGACAACACAGAGCACAATCACTGAAAATGAGCAACCAAGTGCCTCCCTCG 7080
QY
1141 ATAAACAACCTGCTCCAAACGGGAATCCCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1200
Db
7081 ATAAACAACCTGCTCCAAACGGGAATCCCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 7140
QY
1201 GGCCCGCCACAAACGGCACCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1260
Db
7141 GGCCCGCCACAAACGGCACCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 7200
QY
1261 CCCAGCTCGACTGCAACACATCTGTATATTTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
Db
7201 CCCAGCTCGACTGCAACACATCTGTATATTTTTCAGAGAAAGCGATCGATTTCTGGAGG 7260
QY
1321 GAAGTGATATATTCGGTTTTAGATGGTTTAAATACTGAAATGATTTGATTTGATCCA 1380
Db
7261 GAAGGCGACATGTTCCCTTTTCTGGATGGGTTTAAATAATGCTTCCAATGATTTTGACCCA 7320
QY
1381 ATCCCAACACAGAAACAACTTTGATGAATCTCCAGCTTTTAACTTCAACTAATGAG 1440
Db
7321 GTTCCAAATACAAACAACTTTGATGAATCTCTAGTTCTGGTGGCTCGCTCGAGGAA 7380
QY
1441 GAACCAACACATCCCCGGAATACAGTTTAACTTTCTTATTTTCTGATAAAAAATGGA 1500
Db
7381 GATCAACATGCTCCCCCAATATTAGTTTAACTTTTATTTTCTTATTTTCTTAAATAATGAG 7440
QY
1501 GATACGCTACTCTGGGGAAAAACGGAATGATTTGATGAGTGGAGTTTGGAGT 1560
Db
7441 AACACTGCTACTCTGGAGAAAAATGAGATGATTTGATGAGTGAAGATTTGGAGC 7500
QY
1561 GTGAGGAGGACGATTTGGCGCAGGGCTTACGTGATACCAATTTTGGCCCTCGAATC 1620
Db
7501 GTTCAGGAGGATGACCTGGCCGAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGAAT 7560
QY
1621 GAAGGACTCTATACGCGGTTTTTAATCAAAAAATCAGAACAATTTAGTTTGTAGGTTGAGG 1680
Db
7561 GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAACAAATTTGCTGCGAGTTGAGG 7620
QY
1681 CCGTTAGCTTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTAAAGGGTCAACCCAGGAA 1740
|||

Db 7621 CGCTAGCCATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACCTGAGGAA 7680
QY 1741 AGGACATTTTCTTAATCAATAGCATGCAATGACATTTTGGCTTACGAGGTGGGGCGGA 1800
Db
7681 AGAATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGAGGA 7740
QY 1801 ACATGCAAGGTCTAGGACCTGATTTGCAATAGAGATGATCTATCTAAAAATATC 1860
Db
7741 ACATGCAAGGTCTAGGACCTGATTTGCAATAGAGATGATCTATCTAAAAATATC 7800
QY 1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920
Db
7801 TCAGAGCAATTTGACCAATTTAAAGAGGACGACAAAGAGGAGGACTGGTTGGGGCTG 7860
QY 1921 GGTGGCAATGGTGACATCTGACTGGGGTCTTCACCAATTTGGGCATCTGCTACTA 1980
Db
7861 GGTGGTAAATGGTGACATCTGACTGGGGTCTTCCTAACTTGGGCAATTTGGCTACTA 7920
QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db
7921 TTATCCATAGCTGTTCTGATTTGCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7980
QY 2041 GGATGA 2046
Db 7981 GGATAA 7986

RESULT 5
AY430365
LOCUS Lake Victoria marburgvirus strain pp3 guinea pig lethal variant,
DEFINITION complete genome.
ACCESSION AY430365
VERSION AY430365.1 GI:40388387
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 19113)
AUTHORS Hevey M., Negley D., Geisbert J., Jahrling P. and Schmaljohn A.
TITLE Antigenicity and vaccine potential of Marburg virus glycoprotein
expressed by baculovirus recombinants
JOURNAL Virology 239 (1), 206-216 (1997)
MEDLINE 98087840
PUBMED 9426460
REFERENCE 2 (bases 1 to 19113)
AUTHORS Chain P.S.G., Malfatti S.A., Hajjaj A., Vergez L.M., Do L.H.,
Smith K.L. and McCreedy P.M.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore
National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
REFERENCE 3 (bases 1 to 19113)
AUTHORS Ichou M.A., Paragas J., Jahrling P.B., Ibrahim M.S., Loftis L.,
Hevey M. and Schmaljohn A.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort
Detrick, MD 21702, USA
FEATURES
source
1..19113
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="pp3 guinea pig lethal variant"
/isolate="Musoke Kenya, 1980"
/db_xref="taxon:11269"
104..2191
/gene="NP"
104..2191
/gene="NP"
/note="nucleoprotein"
/codon_start=1
/product="NP protein"
/protein_id="AAR85460.1"

Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;			
QY	1	ATGAAGACACATGTTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	60
Db	5941	ATGAAGACACATGTTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	6000
QY	61	ATTTTATGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
Db	6001	ATTTTATGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	6060
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT	180
Db	6061	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT	6120
QY	181	GATTCCTCTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
Db	6121	GATTCCTCTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	6180
QY	241	GTGTAGTACACAGAGGGGAGGAAGCCAAAAATCATCTACAATATAGTGTAAACGGATCCC	300
Db	6181	GTGTAGTACACAGAGGGGAGGAAGCCAAAAATCATCTACAATATAGTGTAAACGGATCCC	6240
QY	301	TCTGAAAAATCCTTCTGTGTAGATCCTCTACCAATCCCGTACATCCGNAATGCAAA	360
Db	6241	TCTGAAAAATCCTTCTGTGTAGATCCTCTACCAATCCCGTACATCCGNAATGCAAA	6300
QY	361	ACTATCCCATCATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
Db	6301	ACTATCCCATCATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	6360
QY	421	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGATACCGAGGCAAAAGTCTTC	480
Db	6361	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGATACCGAGGCAAAAGTCTTC	6420
QY	481	ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGATGCAAAAATGATTTCTCG	540
Db	6421	ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGATGCAAAAATGATTTCTCG	6480
QY	541	CGGCAAGGACAGGGTACCGTCATATGAATCTGACTTCTACTAATAAATATTTGGACAAGT	600
Db	6481	CGGCAAGGACAGGGTACCGTCATATGAATCTGACTTCTACTAATAAATATTTGGACAAGT	6540
QY	601	AGTAAACGGAAACCAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT	660
Db	6541	AGTAAACGGAAACCAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT	6600
QY	661	ACAAAGAACCAAAACATGTGCTCGTCCAAAATACCTCCACACTGCCACAGCCCGTCGG	720
Db	6601	ACAAAGAACCAAAACATGTGCTCGTCCAAAATACCTCCACACTGCCACAGCCCGTCGG	6660
QY	721	GAGTCAAACTCACAGACACCCCAACTGATGCCACCAAACTCAATACACGGACCCCAAGC	780
Db	6661	GAGTCAAACTCACAGACACCCCAACTGATGCCACCAAACTCAATACACGGACCCCAAGC	6720
QY	781	AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	840
Db	6721	AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	6780
QY	841	ACTTCTGATGGGTACCAAGCAAGGGCTTTTCAATCAATGCCACCACTCCCTCACA	900
Db	6781	ACTTCTGATGGGTACCAAGCAAGGGCTTTTCAATCAATGCCACCACTCCCTCACA	6840
QY	901	CAACCAAGCACCCACAGCAAGGAGAAACACAAACCAATTCCTCCAGATGCTGTGACT	960
Db	6841	CAACCAAGCACCCACAGCAAGGAGAAACACAAACCAATTCCTCCAGATGCTGTGACT	6900
QY	961	GAACTAGACAAAAATAACAACATGTCACAAACCGTCCATGCCCCCTCATAACTACCACA	1020
Db	6901	GAACTAGACAAAAATAACAACATGTCACAAACCGTCCATGCCCCCTCATAACTACCACA	6960
QY	1021	ATCTCTACTAAACACACCTTCCAAAACAAACATTTAGCACTCTCTCTGCAACATTAACAAAC	1080
Db	6961	ATCTCTACTAAACACACCTTCCAAAACAAACATTTAGCACTCTCTCTGCAACATTAACAAAC	7020

RESULT 6
AX717721
LOCUS

AX717721

2051 bp

DNA

linear

PAT 15-APR-2003

QY	1081	ACCACCAATGACACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG	1140
Db	7021	ACCACCAATGACACACACAGAGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG	7080
QY	1141	ATAACAACCTTGCCTCCAAACGGGAAATCCACACAGCAAAAGAGCAGCAGCAGCAAAAA	1200
Db	7081	ATAACAACCTTGCCTCCAAACGGGAAATCCACACAGCAAAAGAGCAGCAGCAGCAAAAA	7140
QY	1201	GGCCCCGACCAACGGGACCAACACAGCAAAATGAGCAATTTCCAGCTCTCCCCCCAC	1260
Db	7141	GGCCCCGACCAACGGGACCAACACAGCAAAATGAGCAATTTCCAGCTCTCCCCCCAC	7200
QY	1261	CCCAGCTCGACTGCAACAACATCTTTGTATATTTCAGAGAAAGCGATCGATTTCTGGA	1320
Db	7201	CCCAGCTCGACTGCAACAACATCTTTGTATATTTCAGAGAAAGCGATCGATTTCTGGA	7260
QY	1321	GAAAGTGATATATTCCCGTTTTAGATGGGTTAATAAATCTGAAAATTTGATTTGATCCA	1380
Db	7261	GAAAGTGATATATTCCCGTTTTAGATGGGTTAATAAATCTGAAAATTTGATTTGATCCA	7320
QY	1381	ATCCCAACACAGAAAAAATCTTTGATGAATCTCCAGCTTTAAATACCTTCAACTAATGAG	1440
Db	7321	ATCCCAACACAGAAAAAATCTTTGATGAATCTCCAGCTTTAAATACCTTCAACTAATGAG	7380
QY	1441	GAAACAACACATCCCCCGGAATATCAGTTTAACTTTCTCTATTTTCTGATAAAAAATGGA	1500
Db	7381	GATCAACATGCTCCCCCAATATTAGTTTAACTTTCTCTATTTTCTCTAATAATAATGAG	7440
QY	1501	GATCTGCTACTCTCTGGGGAAAAACGAGAAATGATTTGTGATGACAGAGTTGAGAGTTGGAGT	1560
Db	7441	AACACTGCTACTCTCTGGAGAAAAATGAGAAATGATTTGTGATGACAGAGTTAAGAAATTTGGAGC	7500
QY	1561	GTGACAGGAGGAGATTTGGCGCAGGGCTTACGTGATACCAATTTTGTGGCCCTCGAATC	1620
Db	7501	GTTCAGGAGGAGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAT	7560
QY	1621	GAAAGACTCTACTCTGCGGTTTTAATCAAAAAATCAGAAACAATTTAGTTTGTAGTTGAGG	1680
Db	7561	GAAAGACTTTACACTGCTGCTGTTTTAATTAANAATCAAAACAATTTGCTGTCAGGTTGAGG	7620
QY	1681	CGCTTAGCTAATCAAACTGCTTAAATCTTGGAGCTCTTTGTAAGGGTCACAAACGAGGAA	1740
Db	7621	CGCTTAGCTAATCAAACTGCTTAAATCTTGGAGCTCTTTGTAAGGGTCACAAACGAGGAA	7680
QY	1741	AGGACATTTTCTTAACTCAATAGCATCAATTCGCTTTTGTCTTACGAGGTGGGGCGGA	1800
Db	7681	AGGACATTTTCTTAACTCAATAGCATCAATTCGCTTTTGTCTTACGAGGTGGGGCGGA	7740
QY	1801	ACATGCAAGGTGCTAGGACCTGATTTGTTGCATAGGAATAGGAATCTATCTAAAAAATATC	1860
Db	7741	ACATGCAAGGTGCTAGGACCTGATTTGTTGCATAGGAATAGGAATCTATCTAAAAAATAT	7800
QY	1861	TCAGAAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAAACCTGGCTGGGGCTTA	1920
Db	7801	TCAGAAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAAACCTGGCTGGGGCTTA	7860
QY	1921	GGTGGCAAAATGAGCAATCTGACTGGGGTGTCTCACCATAATTTGGGATCTCTGCTACTTA	1980
Db	7861	GGTGGCAAAATGAGCAATCTGACTGGGGTGTCTCACCATAATTTGGGATCTCTGCTACTTA	7920
QY	1981	TTATCTATAGCTGTTCTGATTTGCTCTGCTGATCTCTGCTGATCTCTCACTAAATACATTT	2040
Db	7921	TTATCTATAGCTGTTCTGATTTGCTCTGATCTCTGCTGATCTCTCACTAAATACATTT	7980
QY	2041	GGATGA 2046	
Db	7981	GGATAA 7986	

DEFINITION Sequence 9 from Patent WO02079239.
ACCESSION AX717721 GI:29890729
VERSION AX717721.1
KEYWORDS Marburg virus (strain Musoke)
SOURCE Marburg virus (strain Musoke)
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Flaviviridae; Marburg-like viruses.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 9 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source location/Qualifiers
1..2051
/organism="Marburg virus (strain Musoke)"
/mol_type="unassigned DNA"
/db_xref="taxon:3372"
/note="chimeric molecule between Marburg virus
glycoprotein 1 and Marburg virus glycoprotein 2"
ORIGIN
Query Match 88.6%; Score 1813.6; DB 6; Length 2051;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 124; Indels 12; Gaps 1;
QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCC 60
DB 10 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCC 69
QY 61 ATTTTAGAGATAGTAGTAATTAATCAACCCCAAAATGGATTTCGGTATGTCCTCGGA 120
DB 70 ATTTTAGAGATAGTAGTAATTAATCAACCCCAAAATGGATTTCGGTATGTCCTCGGA 129
QY 121 CTCGAGAGACAGAGACGTCTCATCTGATGGATTACACTGAGTGGGCAAAAAGTGTCT 180
DB 130 CTCGAGAGACAGAGACGTCTCATCTGATGGATTACACTGAGTGGGCAAAAAGTGTCT 189
QY 181 GATTCCTCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGA 240
DB 190 GATTCCTCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGA 249
QY 241 GTTGTAGTACAGAGGGGAGAGCGAAACATGCTACATATATAGTGTACAGGATCCC 300
DB 250 GTTGTAGTACAGAGGGGAGAGCGAAACATGCTACATATATAGTGTACAGGATCCC 309
QY 301 TCTGGAATCTCTGTTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA 360
DB 310 TCTGGAATCTCTGTTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA 369
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB 370 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 429
QY 421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAAATGTAACGAGGCAAAATCTTC 480
DB 430 GGAGCATTTTTCTGTATGATCGCAT-----ATGTACCGAGGCAAAATCTTC 477
QY 481 ACTGAAGGACATAGCAGTATGTTGTAATAGACAGTGCACAAATGATTTTCTCG 540
DB 478 ACTGAAGGACATAGCAGTATGTTGTAATAGACAGTGCACAAATGATTTTCTCG 537
QY 541 CGGCAAGGACAAAGGTACCGTATATGATCTGACTTCTTACTTAATAAATATTGGACAAGT 600
DB 538 CGGCAAGGACAAAGGTACCGTATATGATCTGACTTCTTACTTAATAAATATTGGACAAGT 597
QY 601 AGTAACGGAACGCAACGAAATGACACTGATGTTTCGGGCTCTTCAAGAAATACAAATCT 660
DB 598 AGTAACGGAACGCAACGAAATGACACTGATGTTTCGGGCTCTTCAAGAAATACAAATCT 657
QY 661 ACAAGAACCAACATGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 720
DB 658 ACAAGAACCAACATGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 717

QY 721 GAGATCAAACTCAAGCAACCCCAACTGATGCCCAAACTCAATACCAAGGACCAAGC 780
DB 718 GAGATCAAACTCAAGCAACCCCAACTGATGCCCAAACTCAATACCAAGGACCAAGC 777
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGGACGAGACGAGACCCACACA 840
DB 778 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGGACGAGACGAGACCCACACA 837
QY 841 ACTTCTGATGCGGTCAACCAAGGAGGCTTTTCATCAACAAATGCCACCCACTCCCTCACCA 900
DB 838 ACTTCTGATGCGGTCAACCAAGGAGGCTTTTCATCAACAAATGCCACCCACTCCCTCACCA 897
QY 901 CAACCAAGCAGCCACAGCAAGAGGAGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 960
DB 898 CAACCAAGCAGCCACAGCAAGAGGAGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 957
QY 961 GAACTAGACAAATAATACAACTGCAACCGTCCATGCCCCCTCATTAACACTTACCACA 1020
DB 958 GAACTAGACAAATAATACAACTGCAACCGTCCATGCCCCCTCATTAACACTTACCACA 1017
QY 1021 ATCTCTACTAAACACCTCCAAACAACTTTTCAGCACTCTCTCTGCACCAATTAACAAAC 1080
DB 1018 ATCTCTACTAAACACCTCCAAACAACTTTTCAGCACTCTCTCTGCACCAATTAACAAAC 1077
QY 1081 ACCACCAATGACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCTCCG 1140
DB 1078 ACCACCAATGACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCTCCG 1137
QY 1141 ATACCAACCTTCCTCCAAACGGGAATCCCAACAGCAAGCAAGAGACGAGCAAGAAAA 1200
DB 1138 ATACCAACCTTCCTCCAAACGGGAATCCCAACAGCAAGCAAGAGACGAGCAAGAAAA 1197
QY 1201 GGCCCGGACACAGGACACCAACCAAGCAAAATGAGCAATTTTCACAGTCTCTCCCTCCAC 1260
DB 1198 GGCCCGGACACAGGACACCAACAGCAAAATGAGCAATTTTCACAGTCTCTCCCTCCAC 1257
QY 1261 CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAGAAAGAAAGCGATCGATTTTCGAAA 1320
DB 1258 CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAGAAAGAAAGCGATCGATCTCTCGAG 1317
QY 1321 GAAAGTGATATATTTCCCGTTTTAGATGGGTAAATACTGAATTTGATTTTGTATCCA 1380
DB 1318 GAAAGCGCATGTTTCCCTTTTTCGATGGGTAAATACTGAATTTGATTTTGTATCCA 1377
QY 1381 ATCCCAACACAGAAACAACTCTTTGTATATTTTCAGAAAGAAAGCGATCGATTTTCGAAA 1440
DB 1378 GTTCCAAATACAAACAACTCTTTGTATATTTTCAGAAAGAAAGCGATCGATTTTCGAAA 1437
QY 1441 GAACAAACACTCCCTCCGAATATCAGTTTAACTTTTCTTATTTTCTGTATAAAATGGA 1500
DB 1438 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATAAATGAG 1497
QY 1501 GATCTGCTACTCTGGGGAACACGAGATGATGTTGATGACAGATTCAGGATTTGGAGT 1560
DB 1498 AACACTGCTACTCTCGGAGAAATGAGATGATGTTGATGACAGATTTAAGAAATTTGGAGC 1557
QY 1561 GTGACGAGGAGCAGATTTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
DB 1558 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTGGCCCTCGAAT 1617
QY 1621 GAAGGACTTATCTGCGGTTTTTAATCAAAATACAGAAACAAATTTAGTTTGTAGGTGAGG 1680
DB 1618 GAAGGACTTATCTGCTGTTTTTAATTAATTAATTAATTAATTTGGTCTGCGAGTTGAGG 1677
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTTTGTTAAGGCTCACACCGAGAA 1740
DB 1678 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTTTGTTAAGGCTCACACCGAGAA 1737
QY 1741 AGGACATTTCTCTTAATCAATAGGATCAATGATGATTTTGTCTTACCGAGGTGGCGCGA 1800
DB 1738 AGAATTTCTCTCTTAATCAATAGATGATGATTTGATTTTCTACTCACAGATGGGAGGA 1797

```
QY 1801 ACATGCAAGGCTCTAGGACCTGATTGTTGCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 1798 ACATGCAAGGCTCTAGGACCTGATTGTTGCATCGGATAGAGACTTGTCCAAAAATATT 1857
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGGAGGAACCTGGCTGGGCTCTA 1920
Db 1858 TCAGAGCAAAATGACCAAAATTAAGGACGACAAAGAGGAGGACTGGTGGGGCTGTG 1917
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTCACCAGATTTGGGCATCTGCTACTA 1980
Db 1918 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTACTTGGGCATTTGCTACTA 1977
QY 1981 TTATCTATAGCTGTCTGATTGCTCTGCTCTGATCTGTGCTGATCTTCACTAAATACAT 2040
Db 1978 TTATCCATAGCTGTCTGATTGCTCTATCTCTGATTTGCTGATCTTTACTAAATATATC 2037
QY 2041 GGATGA 2046
Db 2038 GGATAA 2043

RESULT 7
AF005735      2046 bp      RNA      linear      VRL 23-JAN-1998
Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein
precursor (GP) gene, complete cds.
ACCESSION    AF005735
VERSION      AF005735.1 GI:2459879
KEYWORDS
SOURCE
ORGANISM
Lake Victoria marburgvirus
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE    1 (bases 1 to 2046)
AUTHORS      Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
TITLE        Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL      Virology 240 (1), 138-146 (1998)
MEDLINE      98110148
PUBMED       9448698
REFERENCE    2 (bases 1 to 2046)
AUTHORS      Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE        Direct Submission
JOURNAL      Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
source
1. .2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Germany/Marburg/1967/Ratayczak"
/db_xref="taxon:11269"
1. .2046
/gene="GP"
/gene="GP"
/notes="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40460.1"
/db_xref="GI:2459880"
/translation="MKTTCFLFSLILIOGKTLPILEIASNNPONVDSVCGSLTQKT
EDVHLMGFTLSQKVADSPLEASKEWAFRTGVPKKNVEYTEGEAKCTCYNISVTDPSG
KSLLLDPNINIDYPKCTIHHIQCNPHAQSLHLNGAPFLYDRISTWYGRVF
TEGNTAAMIVNKHMFIRSQGRHNLFTSTNKYWTSSNGTQTNIDTCFGLQDEY
NSTKNQCAPSKISPLTPARPEIKPTSTPTDATTNTDNNDEDLITSGSGSEQ
EPYTTSDAVTKQGLSTMPTSPQSPQPGNNTHDSQGVTEPNTKNTTAQPSMP
PHNTTAISNTSKNFTLSVSLQNTNYDQSTATENEOTSAPSKTLLPTGNLT
AKSTNTKPTPTAAMNTNGHLLTSPSPNPTTQHLVFRKRSILWREGDMFPLDG
LINAIDFPDVENTKTIIDESSSSGASAEEDQASPNISLTILSYFPNINENTAYSGEN
ENDCDAEUKIWSQDDLLAAGLUSWIPFGPIEGLYTAGLKNQNNLVCRLLRANQT
AKSLELLRVLVTTTEERTFSLINRHAIIDFLLTRMGGTCVKVLGPDCCIGIEDLSRNTSEQI
```

```
DOIKDQKEGTGMLGGKWTSDWGLVTLNLGILLISIAVLIALSCICRIPTKYG"
1. .81
/gene="Gp"
mat_peptide
82. .2043
/gene="Gp"
/product="glycoprotein"
/misc_feature
1930. .2019
/gene="Gp"
/notes="encodes transmembrane anchor peptide"
ORIGIN
Query Match      82.9%; Score 1695.6; DB 14; Length 2046;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 219; Indels 0; Gaps 0;
QY 1 ATGAAGACCCATGTTCTCTTATCAGTCTTATCTTAATCAAGGACAAAAAATCTCCCC 60
Db 1 ATGAAGACCCATGTTCTCTTATCAGTCTTATCTTAATCAAGGATAAAAAATCTCCCT 60
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTGGTATGTCGGAAT 120
Db 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTGGTATGTCGGAAT 120
QY 121 CTCAGAGACAGAGAGCTTCATCTGATGGATTACACTGAGTGGGCAAAAGTTGCT 180
Db 121 CTCAGAGACAGAGAGATGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGGCATCCAAGCGATGGCTTTCCAGACAGAGGTGTACTCTCCCAAGAT 240
Db 181 GATTCCCTTTGGAGGCATCCAAGCGATGGCTTTCCAGACAGAGGTGTACTCTCCCAAGAT 240
QY 241 GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACAATATAAGTGTACCGATCCC 300
Db 241 GTTGAGTATACAGAGGGGGAGAGCCAAACATGCTACAATATAAGTGTACCGATCCC 300
QY 301 TCTGAAAAATCTTGTCTTTAGATCTCTTACCACATCCGTGACTATCCGNAATGCAA 360
Db 301 TCTGAAAAATCTTGTCTTTAGATCTCTTACCACATCCGTGACTATCTTAAATGCAA 360
QY 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCTTCATTTATGG 420
QY 421 GGAGCATTTTTCTGTATGATCGCATTTCCCTCCACAAATGTACCGAGGCAAGTCTTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCATTTCCCTCCACAAATGTACCGAGGCAAGTCTTC 480
QY 481 ACTGAAGGGAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
QY 541 CGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGACAAGT 600
Db 541 AGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGACAAGT 600
QY 601 AGTAACGGGAACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTTCAAGAATACAAATCT 660
Db 601 AACAAATGGAACACAAAACGAATGACACTGGATGTTTGGCGCTCTTTCAAGAATACAAATCT 660
QY 661 ACAAGAAACCAAAACATGTGCTCCGTCGAAAATACCTCACCACATCGCCACAGCCCGTCGG 720
Db 661 ACGAAGAAATCAAAACATGTGCTCCGTCGAAAATACCTCACCACATCGCCACAGCCCGTCGA 720
QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCGACCAAACTCAATACCAAGGACCAAGC 780
Db 721 GAGATCAAAACCAAAAGCACCCCAACTGATGCGACCACTCACTCAACACAGACGCCAAGC 780
QY 781 AGTGTATGATGAGGACCTCCGCAACATCCGGCTCAGGGTCCGGAGAACAGGAAACCCACACA 840
Db 781 AATGATGATGAGGACCTCATAACATCCGGTTCAGGGTCCGGAGAACAGGAAACCCATATACA 840
QY 841 ACTTCTGATGGGTGACCAAGCAAGGCTTTTCATCAATGCGCCACCACTCCCTCACCA 900
|||||
```

Db 841 ACTTCAGATGCGGTCACTAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Qy 901 CAAACAGACGCCCAACAGCAGGAGGAAACAACAACCAATCCCAAGATGCTGTGACT 960
Db 901 CAAACAGCAGCCCAACAGCAGGAGGAAACAACAACCAATCCCAAGATGCTGTGACT 960
Qy 961 GAACTAGACAAAATAACACAACTGCACAAACCGTCCATGCCGCCCTCATACACACTACCA 1020
Db 961 GAAACCAACAAACCAACCAACTGCACAAACCGTCCATGCCGCCCTCATACACACTACCA 1020
Qy 1021 ATCTCTACTAACACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCAATTAACAAAC 1080
Db 1021 ATCTCTACTAACACACCTCCAAAGAACAACTTTCAGCACTCTCTCTGATCACTACAAAC 1080
Qy 1081 ACCACAAATGACACACAGAGGACAAATCACTGAAATGAGCAAAACAGTGCCTCCCTCG 1140
Db 1081 ACCACAAATGAGACACAGAGGACAACTGAAATGAGCAAAACAGTGCCTCCCTCG 1140
Qy 1141 ATAACAACCTGCTCCCAACGGGAATCCCAACAGCAAGAGACACAGCAGCAAAAC 1200
Db 1141 AAAACAACCTGCTCCCAACAGGAATCTTACAGCAAGAGACACAAACAGCAAGAA 1200
Qy 1201 GCGCCCGCACAAACGGCACCAACAGCAATGAGCAATTTCCACAGTCCCTCCCTCCAC 1260
Db 1201 GCGCCCGCACAAACGGCACCAATGAGCAATTTCCACAGTCCCTCCCTCCAC 1260
Qy 1261 CCCAGCTGACGACACAACTCTGTATTTATTCAGAAAGAGCGATTCGATTTTCGAAA 1320
Db 1261 CCCAACCCGACACAACTCTGTATTTATTCAGAAAGAGCGATTCCTCTGGAGG 1320
Qy 1321 GAAGTGATATATTCCTGTTTATGATGGTGTAAATATGATGATGATGATGATGATGAT 1380
Db 1321 GAAGTGATATATTCCTGTTTATGATGGTGTAAATATGATGATGATGATGATGATGAT 1380
Qy 1381 ATCCCAACACAGAAACAACTCTTATGATGAATCTCCAGCTTTAATATCTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAGACAGCTCTTATGATGAATCTTATGATGATGATGATGATGATGAT 1440
Qy 1441 GAAACACACATCCCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 GATCAACATGCTCCCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Qy 1501 GATAGCTGCTACTCTGGGGAACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 AACATGCTACTCTGGGGAACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Qy 1561 GTGAGGAGGAGGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGGCGCTGGAATC 1620
Db 1561 GTTACAGGAGGATGACTGGCAGCAGGCTCAGTGGATACCGTTTTTGGGCGCTGGAATC 1620
Qy 1621 GAAGGACTTATGCTGCTGTTTAAATTAACCAACAACTTTGCTGCGAGGTTGAGG 1680
Db 1621 GAAGGACTTATGCTGCTGTTTAAATTAACCAACAACTTTGCTGCGAGGTTGAGG 1680
Qy 1681 CGCTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGAAGGTCACAAACGAGAA 1740
Db 1681 CGCTAGGCAATCAAACTGCTAAATCTTGGAGCTCTTGAAGGTCACAAACGAGAA 1740
Qy 1741 AGGACATTTCTTCAATCAATAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 AGGACATTTCTTCAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy 1801 ACATCAAGGCTGCTAGGACCTGATTTGTCATGAGGATGAGATCTATCTAAATAATATC 1860
Db 1801 ACATCAAGGCTGCTAGGACCTGATTTGTCATGAGGATGAGATCTTCTCAGGAATAT 1860
Qy 1861 TCAGAACAAATGACAAATGAGAGGATGAGAGGAGGAACTGCTGGGGCTCTA 1920
Db 1861 TCAGAACAAATGAGAGGATGAGAGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Qy 1921 GGTGCAATGCTGACATCTGATGCTGGGGTGTCTCACCAATTTGGGGATCTGCTACTA 1980
Db 1921 GGTGTAATGCTGACATCTGATGCTGGGGTGTCTTACTAATCTGGGCAATTTGCTACTA 1980

RESULT 8

MAVSPAA
LOCUS Marburg Virus genomic RNA of GP gene. 2948 bp RNA linear VRL 22-JAN-1996
DEFINITION X68493
ACCESSION X68493.1 GI:296960
VERSION gp gene; structural protein.
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2948)
AUTHORS Bukreyev, A., Volchkov, V.E., Blinov, V.M. and Netesov, S.V.
TITLE The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins
JOURNAL FEBS Lett. 323 (1-2), 183-187 (1993)
MEDLINE 93265932
PUBMED 8495737
REFERENCE 2 (bases 1 to 2948)
AUTHORS Bukreyev, A.A., Volchkov, V.E., Blinov, V.M., Dryga, S.A. and Netesov, S.V.
TITLE The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain
JOURNAL Arch. Virol. 140 (9), 1589-1600 (1995)
MEDLINE 96028047
PUBMED 7487490
REFERENCE 3 (bases 1 to 2948)
AUTHORS Bukreyev, A.A.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1992) A.A. Bukreyev, All-Union Inst. of Molecular Biology, Location/Qualifiers
FEATURES source
1..2948 /organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="Popp"
/db_xref="taxon:11269"
6..2851 /gene="GP"
(6..17)..(2841..2851) /gene="GP"
122..2167 /gene="GP"
/codon_start=1
/product="structural protein"
/protein_id="CAA48507.1"
/db_xref="GI:296961"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
/translation="MKTKLPISLILQIKLPILEIASNNQPNVDVSCGTLQKT EDVHLMGFTLGGQKVDSPLEAKRWAFRTGPPPNVEYTGSEAKTCYNISVDPG KSLLLDPPTNIRDYPKCTIHHIQGNPHAGIALHLWGAFLYDRIAFTMYGRVFE TEGNIAAMIVNKTVMKMFSGQGYRHNLTSTNKTWTSNNGTQNDTTCGALQEV NSTKNQTCAPSKI PSLPTARPEIKPTSTPTDATTNTDNNDDDLITSGSGSEQ EPYTTSDAVTKQGLSSTMPPTSPQSTPQEGNNTDHSQCTVTEPNKNTNTAOPSM PPHNTAISTNNTSKNFSTLSVSLQNTNYDTQSTATENEQTSAPSKTLLPTGNLT AKSTNNTKGPPTTAPNNTNGLTSPSTPNPTQHLVYFRKRSILMREGDNFPDGL LINAIPDFVPVNTKTI FDESSSSGASABEDQASPNISLTSYFNFINENATYSGE ENDCDAELRWVQEDDLAAGLSWIPFGPIEGLYTAGLIGNKNNLVCRRLRLANQT AKSLELLRVTTTEERTFSLINRHAIDFLTWGCTCKVLGPDCCIGIBDLRNTSEQI DQIKKDSQKSGTGWLGKWTSDWGLVTNLGILLLLSLIAVLIALSCIRIFTKYIG"

ORIGIN

Query Match 82.7%; Score 1692.4; DB 14; Length 2948;			
Best Local Similarity 89.2%; Pred. No. 0;			
Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;			
QY	1	ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCC	60
DB	122	ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCC	181
QY	61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGCGATTCGGTATGCTCCGGA	120
DB	182	ATTTTAGAGATAGCTAGTAACATCAACCCCAAAATGCGATTCGGTATGCTCCGGA	241
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTTG	180
DB	242	CTCCAGAGACAGAGAGTCTCATCTGATGGGATTCACATGAGTGGGCAAAAGTTG	301
QY	181	GATTCCTCTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCGA	240
DB	302	GATTCCTCTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCGA	361
QY	241	GTGTAGTACACAGAGGGGAGGAGCAAAACATGCTACAATATAAGTGAACGGATCC	300
DB	362	GTGTAGTATACAGAGGGGAGGAGCAAAACATGCTACATATAAGTGAACGGATCC	421
QY	301	TCGGAAAAATCTTGTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCA	360
DB	422	TCGGAAAAATCTTGTGTAGATCCTCTACCAACATCCGTGACTATCTTAATGCA	481
QY	361	ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTAT	420
DB	482	ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTAT	541
QY	421	GGAGCATTTTCTGTATGATCGCTCCCAACAAATGCTACCGAGGCAAGTCTTC	480
DB	542	GGAGCATTTTCTGTATGATCGCTCCCAACAAATGCTACCGAGGCAAGTCTTC	601
QY	481	ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTT	540
DB	602	ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTT	661
QY	541	CGGCAAGGACAGGATACCGTATATGATCTGACTTCTACTAATAATATTTGGACA	600
DB	662	AGGCAAGGACAGGATACCGTATATGATCTGACTTCTACTAATAATATTTGGACA	721
QY	601	AGTAACGGAAACCAACAGTACACTGATGTTTGGGCGCTCTTCAAGAAATACAA	660
DB	722	ACAAATGGAAACCAACAGTACACTGATGTTTGGGCGCTCTTCAAGAAATACAA	781
QY	661	ACAAAGAACCAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCACAGCCG	720
DB	782	ACAAAGAACCAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCACAGCCG	841
QY	721	GAGATCAAACTCACAAGCACCCCACTGATGCGCAACCACTCAATACCAAGCA	780
DB	842	GAGATCAAACTCACAAGCACCCCACTGATGCGCAACCACTCAATACCAAGCA	901
QY	781	AGTGATGATGAGGACTTCGCAACATCCGCTCAGGTCGGAGAACGAGAACCC	840
DB	902	AGTGATGATGAGGACTTCGCAACATCCGCTCAGGTCGGAGAACGAGAACCC	961
QY	841	ACTTCTGATGCGGTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACT	900
DB	962	ACTTCTGATGCGGTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACT	1021
QY	901	CAACCAAGCACCGGTCACAGCAAGGAAACAAACAAATTCAGAGTCTGTGACT	960
DB	1022	CAACCAAGCACCGGTCACAGCAAGGAAACAAACAAATTCAGAGTCTGTGACT	1081
QY	961	GAACTAGACAAAAATAACCAACTGACCAACCGTCCATCCCTCATATAACTACC	1020
DB	1082	GAACTAGACAAAAATAACCAACTGACCAACCGTCCATCCCTCATATAACTACC	1141
QY	1021	ATCTCTACTAAACAACTCTCCAAACAACTTACGACTCTCTCTGCAACCAAT	1080

RESULT 9

DB	1142	ATCTCTACTAAACACCTCCAAAGAACAACTTTCAGCACCCCTCTCTGTATCATCA	1201
QY	1081	ACCACCAATGACACACACAGAGACACAACTCACTGAAAATGAGCAAAACAGTGC	1140
DB	1202	ACCACCAATGACACACACAGAGACACAACTGAAAATGAGCAAAACAGTGC	1261
QY	1141	ATAACAACTCTGCTCCAAACGGGAAATCCCAACAGCAAGAGACACAGAGCA	1200
DB	1262	ATAACAACTCTGCTCCAAACGGGAAATCCCAACAGCAAGAGACACAGAGCA	1321
QY	1201	GGCCCCGCCACACAGGACCAAAACAGCAAAATGAGCAATTCACCAAGTCTCC	1260
DB	1322	GGCCCCGCCACACAGGACCAAAATGAGCAAAATGAGCAATTCACCAAGTCTCC	1381
QY	1261	CCAGCTCGAGCTGCACAACTCTTGTATATTTTCAGAGAAAGCGATTCGATTT	1320
DB	1382	CCAAACCCGACACACAACTCTTGTATATTTTCAGAGAAAGCGATTCCTCTG	1441
QY	1321	GAAAGTGATATATTTCCCGTTTTAGATGGGTTAAATACTGAAATTTGATTTG	1380
DB	1442	GAAAGTGATATATTTCCCGTTTTAGATGGGTTAAATACTGAAATTTGATTTG	1501
QY	1381	ATCCAAACACAGAAACAACTCTTGTATGAACTCCAGCTTTAATACTCACTAT	1440
DB	1502	GTTCCAAAATACAAAGACGATCTTTGATGAATCTTCTAGTCTCTGGCTGAG	1561
QY	1441	GAAACAAACACTCCCGCGAATATCAGTTTAACTTCTCTTATTTTCTGATAAA	1500
DB	1562	GATCAACATGCTCCCGCAATATCAGTTTAACTTCTCTTATTTTCTGATAAA	1621
QY	1501	GATCTGCTACTCTGGGGAAAAACGAGAAATGATGTGATGACAGAGTTGAGG	1560
DB	1622	AACACTGCTACTCTGGGAAAAACGAGAAATGATGTGATGACAGAGTTAAGA	1681
QY	1561	GTGAGGAGGACGATTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTG	1620
DB	1682	GTTCAGGAGGATGACTCTGGCAGAGGCTCAGTTGGATACCGTTTTTGGCC	1741
QY	1621	GAAAGACTCTACTGCGCGTTTTTAAATCAAAATCAGAACTTAAAGTTTGT	1680
DB	1742	GAAAGACTCTACTGCGCGTTTTTAAATCAAACTTAAAGTTTGTAGGTTGAG	1801
QY	1681	CGCTTAGCTAAATCAAACTGCTAAATCTTGGAGCTCTTTTAAAGGTCAC	1740
DB	1802	CGCTTAGCTAAATCAAACTGCTAAATCTTGGAGCTCTTTTAAAGGTCAC	1861
QY	1741	AGGACATTTTCTTAACTCAATAGGATGCAATTTGCTTTTACGAGGTGGGCG	1800
DB	1862	AGGACATTTTCTTAACTCAATAGGATGCAATTTGCTTTTACGAGGTGGGCG	1921
QY	1801	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTA	1860
DB	1922	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTTCTCA	1981
QY	1861	TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGT	1920
DB	1982	TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGT	2041
QY	1921	GGTGCAAAATGGTGACATCTGACCTGGGGTGTCTCACCAATTTGGGCACT	1980
DB	2042	GGTGCAAAATGGTGACATCTGACCTGGGGTGTCTCACCAATTTGGGCACT	2101
QY	1981	TTATCTATAGCTGTTCTGATTCCTGCTCTGCTGCTGCTGCTGCTGCTGCT	2040
DB	2102	TTATCTATAGCTGTTCTGATTCCTGCTCTGCTGCTGCTGCTGCTGCTGCT	2161
QY	2041	GGATCA 2046	
DB	2162	GGGTAA 2167	

MVVIRPR	LOCUS	2944. .3933	19112 bp	RNA	linear	VRL 22-JAN-1996
DEFINITION	Marburg virus (Popp)	NP, VP35, VP40, GP, VP30, VP24, L genes.				
ACCESSION	229337					
VERSION	229337.1	GI:450908				
KEYWORDS	GP protein; L protein; NP protein; VP24 protein; VP30 protein; VP35 protein; VP40 protein.					
SOURCE	Lake Victoria marburgvirus					
ORGANISM	Lake Victoria marburgvirus					
REFERENCE	1	(bases 2851 to 5818)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.					
TITLE	The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses					
JOURNAL	PEBS Lett. 322 (1), 41-46	(1993)				
MEDLINE	93245956					
PUBMED	8482365					
REFERENCE	2	(bases 5824 to 8669)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.					
TITLE	The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins					
JOURNAL	PEBS Lett. 323 (1-2), 183-187	(1993)				
MEDLINE	93265932					
PUBMED	8495737					
REFERENCE	3	(bases 1 to 19112)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and Netesov,S.V.					
TITLE	The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain					
JOURNAL	Arch. Virol. 140 (9), 1589-1600	(1995)				
MEDLINE	96028047					
PUBMED	7487490					
REFERENCE	4	(bases 1 to 19112)				
AUTHORS	Bukreyev,A.A.					
TITLE	Direct Submission					
JOURNAL	Submitted (13-JAN-1994)					
MEDLINE	Biology, Laboratory of molecular virology, Koltsovo, Novosibirsk region, Russia, 633159					
PUBMED	Location/Qualifiers					
FEATURES	1. .19112					
source	/organism="Lake Victoria marburgvirus"					
	/mol_type="genomic RNA"					
	/strain="Popp"					
	/db_xref="taxon:11269"					
gene	103. .2190					
	/gene="NP"					
CDS	103. .2190					
	/gene="NP"					
	/function="nucleoprotein"					
	/citation=[3]					
	/number=1					
	/codon_start=1					
	/label=NP					
	/product="NP protein"					
	/protein_id="CAA82536.1"					
	/db_xref="GI:450909"					
	/db_xref="GOA:P35263"					
	/db_xref="UniProt/Swiss-Prot:P35263"					
	/translation="MDLSLEIGTKPTAPHVRNKVILFDTNHQVSIQNIIIDAINS GIDGLLEGLLTCLVSHYNSDKDFNTSPIAKYLRDAGYFDFVKNADATRFIDV INPEHYSVIYLNALKTLESTESQRIGLFPCSLPLKLVVGDRASTEKALROVTV HQRGIVTVYPNHLLTGHMKVIIGILRSSFLFKVLVHGVNLVTGHDAVDYIINSV GQTRFSGLLIVKTVLPILQKTDSGVTPLHVRTSKVKNEVASFQALSNLARHGEYA PFARLNLSGNLLEHGLYQLSAIALGVATGATAGVATAGVNVGEQYQQRLEAAHDAE I KLQRHHEBQIIAIEDDEBTLEQFLHQTKEITHSQTLAVLSQREKRLAAAEI NNTVDEGQFQSQVQSFLNDPTFVEVTQARPNRPTALPPVPDKNIEHSTEDS SSSGSFDVLDNDEPALLNDEBTDLDDSVMIPTTSREFQGIAPPPOSDLNNSQKOE DESTNPKFLYQELPLPVQEDDESEYTTDSQESIDQPSDNEQGVDPPLPYAOE KRQDPTQHPAVSQDPFGSIGDVNGDILFIPRSPSPAPQEDTRAREAYELSPDTFN YEDNQNVQFORVYVTKGRFLYLPNDLLQTNPESLITAVEEQNPVSAKELQADWFD MSFDERRHVAMN."					
	2944. .3933					
gene						


```
/number=5
/codon_start=1
/label=VP30
/product="vp30 protein"
/protein_id="CAA82540.1"
/db_xref="GI:450913"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="MQPRGRSRTRNHTQASSIYHETQLPSKPHYTNHHPARMSST
RSSAESPTNHPRARPPPTNLSKPPPKDMCMNMKIGLPCTDPTCNVRDHDLDNLT
NRELLLMARKWLPNTDFTSLQDCGSPSLSKGLSKDKQSQTKDVLTLNLHLNLY
LHRSDIGKDETSRLAASLSTCAGIRKTRNRLDINTWELHINHENLPDQNGVINQTY
TGIHDKGGQFEALMQGWDKRSISLFVQAALYVNNNIPCESSTSVQASVDFHILPOS
QSKGQ"
gene
10205..10966
/gene="VP24"
CDS
10205..10966
/gene="VP24"
/citation={3}
/number=6
/codon_start=1
/label=VP24
/product="vp24 protein"
/protein_id="CAA82541.1"
/db_xref="GI:450914"
/db_xref="UniProt/Swiss-Prot:P41325"
/translation="MAELSTRYNLPANVTEKSIINDLNSLNARWIKEPSVGGWTKWGN
FVPHIPNTGMALLHHLKSNFVVPWQOTRNLPSHLFKNPKSTIIEPPLALRLLGVAL
KDEELQOQSLIPGFRSIVHMLSEMLLEVTSAIHISPNLLGIYLTSPKELMAGVNF
FNKMFILHVNDHGKPSSEIKLTQQQIIITRVNMGFLVEVRRIDIEPCSETVLSES
VVFGLVABAVLRHBSQMEKQGPLDUTQYVNSKIAI"
gene
11479..18474
/gene="L"
CDS
11479..18474
/gene="L"
/function="major component of RNA-dependent
RNA-polymerase"
/citation={3}
/number=7
/codon_start=1
/label=L
/product="L protein"
/protein_id="CAA82542.1"
/db_xref="GI:450915"
/db_xref="GOA:P35262"
/db_xref="UniProt/Swiss-Prot:P35262"
/translation="MQHPTQYDPAKLSPIILLDQCDLLARSLGLYSHYSHNPKNCR
IPHHYLRKNSLTALTKTQNGSILTVPPFHSIWDHILTSIQYDAINHVDVDFKLLPSEL
VKYAWDNFLKAYLNKILGLDHPFASARQWEDFSPKNPYWGMILLVHLSQLAR
RIKQGRSLRNMKEIGTDLFLFIADFIIFKVPVKTIIIRNAVSLQASKPLRVWYRD
QNLTPYLCDDDEFIVSASVECFIMIKDVFIERNTWEICARAWLEDSDGADYLPDLVL
GELYNQGDQIIAMYLEDDGPKLKHLEPLCVSCIOTHGIFTPGKWFQSORIESYYEEL
CSLWKFPLISGNKASCAQNFIKTIIOGKLTPOQYCELSLQKHGHPVLYIDVALDKV
KHAQSVKILKPKWFETFCVFKFIVAKNHTHSQGSWYKTTMDLHLTPLYLRQHIVSNS
FPSQAEIYQHLMEWTFVBEHPLFUSFKIISDLSIFIKDRATAVNQECWDSVDFRSVLGY
```

Query Match 82.7%; Score 1692.4; DB 14; Length 19112;

Best Local Similarity 89.2%; Pred. No. 0;

Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

```
QY 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC 60
DB 5940 ATGAAGACACATGCTCTTTATCAGTCTTATCTTAAATTCAGGGATAAAAACTCTCCCT 5999
QY 61 ATTTTAGAGATAGCTAGTAAATATCAACCCCAAAATGTGGATCGGTATGCTCCGGAAC 120
DB 6000 ATTTTAGAGATAGCTAGTAAATATCAACCCCAAAATGTGGATCGGTATGCTCCGGAAC 6059
QY 121 CTCGAGACAGAGACGCTCCATCTGATGGGATTCACACTCAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCGAGACAGAGAGATGCTCATCTGATGGGATTCACACTCAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
```

```
DB 6120 GATTCCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 6179
QY 241 GTTCAGTACACAGAGGGGAGGAGCCAAAAACATGCTACAATATAAGTGTAAACGATCCC 300
DB 6180 GTTGATATACAGAAGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGATCCC 6239
QY 301 TCTGAAAAATCCTTGTCTGTAGATCCTCTACCAACATCCGTGATATCCGAAATGCAAA 360
DB 6240 TCTGAAAAATCCTTGTCTGTAGATCCTCTACCAACATCCGTGATATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG 420
DB 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGCGCAAGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCCAACAAATGTACCGAGGAGAGTCTTC 6419
QY 481 ACTGAAGGGAAATAGCAGCTATGATGTTCAATTAAGACAGTGCACAAAATGATTTTCTCG 540
DB 6420 ACTGAAGGGAAATAGCAGCTATGATGTTCAATTAAGACAGTGCACAAAATGATTTTCTCG 6479
QY 541 CGGCAGGACAAAGGTACCGTACATATGAAATCTGACTTCTACTAATAAATATTTGGACNAGT 600
DB 6480 AGGCAAGGACAAAGGTACCGTACATATGAAATCTGACTTCTACTAATAAATATTTGGACNAGT 6539
QY 601 AGTAAACGGAAACGAAACCAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 660
DB 6540 AACATGGAACACAAACGAAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 6599
QY 661 ACAAGAAACCAAAACATGTGCTCCGTCMAAAATACCTCCACCACTGCCACACGCCGCTCGG 720
DB 6600 ACGAAGAAATCAAAACATGTGCTCCGTCMAAAATACCTCCACCACTGCCACACGCCGCTCGG 6659
QY 721 GAGATCAAACTCACHAGCACCCCAACTGATGGCCACCAAACTCAATACACGAGACCCAGC 780
DB 6660 GAGATCAAACTCACHAGCACCCCAACTGATGGCCACCAAACTCAATACACGAGACCCAGC 6719
QY 781 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCGGGAGACGAGAACCCACACACA 840
DB 6720 AATGATGATGAGGACCTCATAACTCCGGTTCAGGGTCGGGAGACGAGAACCCATATACA 6779
QY 841 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAATGCCCCACCACTCCCTCAACA 900
DB 6780 ACTTCAGATCGGTCACTAAGCAAGGGCTTTTCATCAACAATGCCCCACCACTCCCTCAACA 6839
QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAACAAACAAACCACTTCCCAAGGTACTGTGACT 960
DB 6840 CAACCAAGCAGCCACAGCAAGGAGGAAACAAACAAACCACTTCCCAAGGTACTGTGACT 6899
QY 961 GAACTAGACAAAAATAACACAACTGSCACACCGTCCATGCCCCCTCATATAACTACTACCACA 1020
DB 6900 GAACCAACAAACAAACCAACAGGACACACGTCATGCCCCCACCACCACTACTGCA 6959
QY 1021 ATCTCTACTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
DB 6960 ATCTCTACTAAACACACCTCCAAAGAACTTCAGCACTCTCTCTGATATCACTACAAAC 7019
QY 1081 ACCACCAATGACACACAGAGGACCAATCACTGAAATGAGCAAAACGATGCCCCCTCG 1140
DB 7020 ACCACCAATGACACACAGAGGACCACTGAAATGAAATGAAACCAACGATGCCCCCTCG 7079
QY 1141 ATAACCAACCTCGCTCCAAACGGAATCCACACAGCAAGAGGACCAACGAGCAGCAAAAAA 1200
DB 7080 AAAACCAACCTCGCTCCAAACGGAATCTTACCAAGCAAGAGGACCTTACAGCAGAAA 7139
QY 1201 GGCCCCGCCACAAACGAGCAGCAACCAACGAAATGAGCACTTTCACCACTCTCCCCCACC 1260
DB 7140 GGCCCCGCCACAAACGAGCAGCAACCAACGAAATGAGCACTTTCACCACTCTCCCCCACC 7199
QY 1261 CCAGCTGCTGACCAACATCTTGTATATTTTCAAGAAAGCGATCGATTTTCTGGAAA 1320
```


Db	7200	CCCAACCCGACACACAACTCTTGTATATATTCAGAAAGAAAAGAAATGATCCTCTCGAGG	7259
Qy	1321	GAAGGTGATATATTCCTGGTTTTAGATGGGTAAATAAATACTGAAATGATTTGATCCA	1380
Db	7260	GAAGGCACATGTTCTCTTCTGGACGGGTAAATAAATGCTCCAATGATTTGATCCA	7319
Qy	1381	ATCCAAACACAGAAACAATCTTTTGATGAATCTCCCGAGCTTTAAATATCTCAACTAATGAG	1440
Db	7320	GTTCCAAATACAAGACGATCTTTGATGAATCTTCTAGTTCGTGCTCGGCTGAGGAA	7379
Qy	1441	GAACAACACACTCCCGCGAATATACAGTTTAACTTTCTCTTATTTCTCTGATAAAAATGGA	1500
Db	7380	GATCAACATGCTCCCCCAATATACAGTTTAACTTTATCTTATTTCTTAATATAATGAA	7439
Qy	1501	GATACTCCCTACTCTGGGGAACAGGAATGATTTGTGATCAGAGTTGAGGATTTGGAGT	1560
Db	7440	AACACTGCCCTACTCTGAGAAATAGAAACGATTTGTGATCAGAGTTAAGAAATTTGGAGC	7499
Qy	1561	GTGACAGACAGATTTTGGCGCAGGGCTTAGCTGGATACCATTTTTTGGCCCTGGAATC	1620
Db	7500	GTTTCAGGAGATGACCTTGGCAGCAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAATC	7559
Qy	1621	GAAGGACTCTATCTGCGCGTTTAAATCAAAAAACAGAACTTAACTTTGTGATGTTGAGG	1680
Db	7560	GAAGGACTTATATCTGCTGTTTAAATTAATAAACCAAAACAAATTTGGTCTGCAGGTTGAGG	7619
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTCGAGCTCTTGTAAAGGGTCACAAACCGAGGAA	1740
Db	7620	CGTACGCCAATCAAACTGCCAATCCTTGGAACTCTTATTAAGATTCACAACCGAGGAA	7679
Qy	1741	AGGACATTTCTTAAATAGCATGCAATTGACTTTTTTGTCTTACGAGGTGGGCGGA	1800
Db	7680	AGGACATTTCTTAAATAGCATGCAATTGACTTTTTTGTCTTACGAGGTGGGCGGA	7739
Qy	1801	ACATGCAAGTCTGAGACCTGATTTGTCATAGGATAGAACTATCTAAAAATATC	1860
Db	7740	ACATGCAAGTCTGAGACCTGATTTGTCATAGGATAGAACTATCTAAAAATATC	7799
Qy	1861	TCGAAACAAATCGACAAATCGAAGAGATGAACAAAGAGGAGAACTGGCTGGGCTCTA	1920
Db	7800	TCGGAACAATTTGNCAAATCAAAAAGATGAACAAAGAGGGGACTGGTGGGCTCTA	7859
Qy	1921	GGTGGCAAAATGGTGACATCTGACGGGTGTTCTCAACCAATTTGGGCACTCTGCTACTA	1980
Db	7860	GGTGGTAAATGGTGACATCCGACGGGTGTTCTTACTTAACCTGGGCACTTTTGGCTACTA	7919
Qy	1981	TTATCTATAGTGTCTGATGCTCTGCTGATCTGCTGATCTCTCAATAATACAT	2040
Db	7920	TTATCCATAGCTGCTTTGATTTGCTATCTATCTGATTTGCTGATCTTTACCAAAATATC	7979
Qy	2041	GGATGA 2046	
Db	7980	GGGTAA 7985	
RESULT 10			
LOCUS	AY358025	19151 bp RNA linear	VRL 27-AUG-2003
DEFINITION	Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin, complete sequence.		
ACCESSION	AY358025		
VERSION	AY358025.2	GI:34304751	
KEYWORDS			
SOURCE	Lake Victoria marburgvirus		
ORGANISM	Lake Victoria marburgvirus		
REFERENCE	1 (bases 1 to 19151)		
AUTHORS	Bowen, M.D., Thurman, K., Minor, E., Ibrahim, M.S., Meyer, R.F., Malfatti, S.A., Do, L.H., Smith, K.L., McCready, P.M. and Chain, P.S.G.		
TITLE	Submitted (01-AUG-2003) Viral Sequencing Group, Lawrence Livermore		
JOURNAL	National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA		

```

KSLLDPPTNVRDYPKCKTIHHIQQONPHAQIALHLWGAFPLVDRIASITWYRGKVP
TEGNTAAAMIYKNTVHKMIFSRQGGIRHNMNLSTNKNYWTSSNGQTNDTCFCGTHHEY
INSTQKCAPSKTPPPPTARPEIKPTPTDTRNLNTNPNSSDDEDTLTSGSGSEQ
ELYTSDAVTQGLSMTPTSPQPOGTPQCGNNTHSQDATTLENTNTTAQPPPT
SHNTTITGNTNKGSLTGLAPPONTNPNQSMATENEKTSAPKATLPTENPTT
EKSTNNTKSPITMBENTTNGHFTSSPTNSTTOHLIYFRKRSILWREGDMFPFLDG
LINAPIDPDPVNTKTIIDESSSSGASTEEDQHASSNLSLTLVYLPHTSNNTASGEN
ENDCAELIKRISWTDLAAGLSWIFFPGPIEGLYTAGLIGNQNNLVGRRRLANQT
AKSLDLVRVTTEERTFSLINRHAIDFLTLRWGCTKVLGPDCCIGIBDLSRNSIQI
DQIKKDEQKEGTGMLGCKWMTSDMGVLTNLGLILLLSIAVLIALSCICRIFTYIG"
8888..9733
/gene="VP30"
CDS
8888..9733
/gene="VP30"
/codon_start=1
/product="VP30 protein"
/db_xref="GI:33868623"
/translation="MQQPRGRSRTNRHQAIPISIYHETQLPSKPNVTNHHPRARMSST
RSSTESSPTNHIPRARPPSTNLSKPPPPPPDCMCRNMKIGLPCDTLTCNRDHDLDNLT
NRELLLMARKMLPNTDVFKSPQDCGSPSLSKGLSKDKQEQTKDVLTLNGLHLNY
LHRSIGKLDSTLRAALSLTACAGIKRTNRSLINTMTLHNENLPDQNGVTKQTY
TGIHDKGQPEALWQGDKRESISLFVQAALYVNNNIPCESSISVQASVDHFILPQS
QKGCQ"
10226..10987
/gene="VP24"
CDS
10226..10987
/gene="VP24"
/codon_start=1
/product="VP24 protein"
/db_xref="GI:33868624"
/translation="NAELSTRYNLPANVTEKSIINDLNSTARWIKEPSVGGTVKMGV
FVFHPTMGNTLHLHKNFVFPWQQRNLFSLFKPNKSTIIEPFLAILRLLGVAL
KDOELQQLSIPGRSIVHMLSEWLLLEVTSIAHISPNIILGLVLTSDMPKILMAGVNF
FNKMTLHVNDHGRPSIEIKLTGQQLITRVNMGFLVEVRRIDIEPCCEGTILSES
VFGLVAAEVRHESQMEKQPLNLTQIMNSKIAI"
11500..18495
/gene="L"
CDS
11500..18495
/gene="L"
/note="major component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="L protein"
/db_xref="GI:33868625"
/translation="MQHPTQVDPARLSSPVLDOCLLARSILGLYSHVSHNPKLRNCR
IPHHIYRARNSTALTKFLQNCSLITVPFHSINWDLITSIQYDAINHVDLDFKLLPSEL
VKIANDNEFLKTYLNLKILRDLHVPFARSOCEDFSKPNPYTMGMLLLVHLSQLAR
RIKQGRSLRNMKFIGDLELFGIADFIIFKVPVKTIIRNAVLSQASPKGLRVMYRD
QNLTPYLCDDDEPISVASYECFIMIKDVFIERVNTWEICARAWLEDSDGADYPPLDVL
GELVYNGDOIIMAYLEDGFKLTKHLEPLCVSCIOYTHGIFTPRKYWFQSMIKSYDEL
CDNLKLRISDNKASCAQNFIKTIIOARLTPOQYCELPFLQKHGHPVLYNDVALDKV
KHAQATKILKPKVMEFCVPKFIIVAKNHHYSQGSWYKTHDHLTPYLRLQHVSN
PFSQAEIYOHLMWEIVFEHEPLFSTKIIDSLSIFIKRATAVNQECWDSVDFRSVLGI
NNPVFRQKRVPELQGDALSLNQIDLEFAKLEYLAPS YRNFPSLKEKELNIGRTF
GKLPYRVNRVOTLAALLADGLAKAPPSNMVMVTEREQEALLHWHHNSASIGEN
ATVRGASFVTLKALAFRYBFRHFNYCNRCYGVKNLFDMMHFLPLCMVNSVDF
VSPHCVTEDNENPPDCANAVHYHLGGIEGLQOKLWTCISCAOITLVELTKLKLS
SVMGONQICITLUSLFPIDAPNDYQNEAELNARVAELAITTSGSLFKPEETFVH
SGFIYFGKQYGLVOLQPSLTKMARCGPLSDSIFDDLOGLSASIGTSFERGTSETH
IPPSWIASFHSMLAINLNQHLFPGLFNIDISCFKPLTTFSEKJALITPQVLGG
LSFLNPEKLVNLDISDPLTSGLFQKNALIEFLGKGLFYLIIAKKPGLADADSVFVNP
LGLNVPAGGREITTFURQTVRENTISQNRINISLPHIGSDLEDOKVCEWLLSNPVM
SRFADIPSRTSGKROVLGVLEGTRTLASRTISLTTEGTMKMLKRELTRNWKSW
PSYDALDDLSLESKEFTCTVDVANFLRAYSWSDVLGKRLIGATLPCLEQPKYKW
INLSDELREQFHSSESELINLLPYDCKEURLGSGHDTNELNVSCALDRKVVQKHP
VKRLAWTIGNRAPYIGSRTEDEKIGYPLRVNCPGSAALKEAIEWVSRLLWVTQGTADRE
KLLIPLNLSRVNLDQTVNLFPYTHYSNVHRYNDQYQHSFMANRMSNTSTRAIIS
TNTLGYAGGQAAIDSNIIIFONTINLGVAVLDIALSLAKLSASNVTFRMLNKNCT
RHVSEYLPFPDLPLDNLKMYMDNELVNDNPLCSGIKRGLGRVSRSTLSILNVSDI
GSDYDPTTAAWTLGRTIIGVSPSDESSQSTDPISSGCTKTFWTHFLVYPVESIPYAFG
ANLIVESLSLRKISIKNLSIDTFLISSTIRNLSHRISUKIQSTFRRHLVLRLAHHI

```

```

PLISMLGSGAGEKSSDAVRLTASVQNFIIHNFSCIMKKGOSLPWLVPSEGGQ
LKPKLKLQRLSDLLSPDKVQRCKTLADTCCSIDSPWYPSKSTRTHYASLAWYRD
KANKVQPTFSLHNLCSFLELSSTSVSSNQVTSKYIVHPEDIPEINTRKLIKEY
GSTALQMDIKSPSEQNLENCRPSKDIFKDNKQITKHQDQYKEKSSQOMSPED
NMOTFLAYHNSSPSOTIKSIDVHEDFASRVILNSKINNENLTDCTINTNLLTPTG
TEFLDTPLOSQSVYSSTPRERSLLSREOASVLYVDCSNIPSIISLDGPRNMSDNOIQ
MLINAYKEDLHACFDSNQFCRPTGVSVSSMYKLYDLAPGCELRAICLAEGESGARL
LLKWKETDHLFFNTLATDSQEAELISGRVIFRMLYINDRISALSERLLINLNUTIQ
ITDITNPLMDSVITQYLPEDSDILTDMAETTKEDTQBLQYKTIWINIWRTPNIPKIS
IIVKVLIDIEGTFLFMKNAIOYQGVQKPKPYSSNAKNSEWLYLCCRKSRQLQDFF
DOVGIFLICAMSRQROAIPYWLKHIEKNIPASLHPEFLTGLFSPSLSSFCRHYTIPF
SEKALPHKQSVYVRQKQHLHLSMLDYENNSPLDLRNHPICSLRGKITKYVNDILK
LNLVIVKAVEKGKWSQVETLPPNHSVCIVHVDHECTGCEKRLLLKLLDFIRNTKIAEQ
KLMNRVIGIILFPFGLSKRSLSRA"
complement(19133..19151)
ORIGIN
Query Match 80.8%; Score 1653.6; DB 14; Length 19151;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 244; Indels 0; Gaps 0;
Qy 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5960 ATGAGGACTACATGCCCTTTTATCAGTCTTATCTTAATTCAGGGATAAAACTCTCCT 6019
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 61 ATTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6020 ATTTGGAGATAGGCAGTAACGATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6079
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 121 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGCGGCAAAAAGTCT 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6080 CTCAGAAAAACAGAAAGATGTCATCTGATGGGATTTACACTGAGCGGGCAGAAAGTCT 6139
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 181 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6140 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACTCTCTTAAGAT 6199
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 241 GTTGAGTACACAGAGGGGAGGAAAGCCAAAACATGCTCAATATAAAGTGTACCGATCCC 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6200 GTTGAGTACAGGAAGGAGGAAAGCCAAAACATGCTCAATATAAAGTGTACCGATCCC 6259
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 301 TCTGGAATACCTTGTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6260 TCTGGAATACCTTGTGTAGATCCTCTCCACCAACGTCCTGACTATCCTAAATGCAAA 6319
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 361 ACTATCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG 420
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6320 ACTATCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG 6379
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAACATGTATCCGAGGCAAGTCTTC 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6380 GGAGCATTTTTCTATATGATCGCATTCCTCCACAAACATGTATCCGAGGCAAGTCTTC 6439
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 481 ACTGAAGGACATAGCAGCTATGATGTCAATAGCAGTGCCACAAATGATTTCTTCG 540
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6440 ACTGAAGGACATAGCAGCTATGATGTCAATAGCAGTGCCACAAATGATTTCTTCG 6499
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 541 CGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAA 600
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6500 AGGCAAGGACAAAGGTCACCGTCATATGAATCTGACTTCTACTATAATAATTTGGACAA 6559
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 601 AGTAAACGGAACGCAAAAGATGACATGATGTTTTCGCGCTCTTCAAGAATACAAATCT 660
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6560 AGCAACGGAACGCAAAAGATGACATGATGTTTTCGCGCTCTTCAAGAATACAAATCT 6619
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 661 ACAAGAACCAACATGTGCTCCGTCCAAAATACCTCCACCAAGTCCGCCACAGCCCGTCCG 720
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6620 ACGAAGAACCAAAAGATGTGCTCCGCTTAAAAACACCCCAACCCGCCACAGCCCGTCCG 6679
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6680 GAATCAAAACCAACAGACCCCAACCGATGCCACTAGACTCAACACCAACCAACCAAC 6739
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


misc_feature		/product="glycoprotein" 1930..2019 /gene="GP" /note="encodes transmembrane anchor peptide"	
ORIGIN			
Query Match	80.8%;	Score 1652.4;	DB 14; Length 2046;
Best Local Similarity	88.0%;	Pred. No. 0;	
Matches 1800;	Conservative 0;	Mismatches 246;	Indels 0; Gaps 0;
QY	1	ATGAAGACACATGTTTCTCTTATCATGCTTATCTTAATTCAGGGACAAAAATCTCCCC	60
Db	1	ATGAGACTACATGCTCTTTATCAGTCTTATCTTAATCCAGGGATAAATACTCTCCCT	60
QY	61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAAT	120
Db	61	ATTTTGGAGATAGCCAGTAAACGATCAACCCCAAAATGTGATTCGGTATGCTCCGGAAT	120
QY	121	CTCCAGAGACAGAGACGTCATCTGATGGGATTCACTGAGTGGGCAAAAGTTGCT	180
Db	121	CTCCAGAAAACAGAAAGATGTCATCTGATGGGATTTACATGAGCGGGCAGAAAGTTGCT	180
QY	181	GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
Db	181	GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCTAAGAT	240
QY	241	GTGTAGTACACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGTAAACGATCCC	300
Db	241	GTGTAGTACACGGAAGGGGAGGACCAAAATGCTACAATATAGTGTACGGATATCCC	300
QY	301	TCGTGAAAATCTCTGCTGTAGATCTCTTACCAACATCCGTGATATCGGAAATGCAAA	360
Db	301	TCGTGAAAATCTCTGCTGTAGATCTCTCCACCAACGTCGTGACTATCTCTAAATGCAAA	360
QY	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG	420
Db	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGCGCAGGGGATCGCCCTTCATTTGTGG	420
QY	421	GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGTCTTC	480
Db	421	GGAGCATTTTTCTGTATGTCGATTCGCTCCAAACATGTACCGAGGCAAGTCTTC	480
QY	481	ACTGAAGGGAAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG	540
Db	481	ACTGAAGGGAAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG	540
QY	541	CGGCAAGGACAGGGTACCGTCAATGATCTGACTTCTACTAATAAATTTGGACAAGT	600
Db	541	AGGCAAGGACAGGGTACCGTCAATGATCTGACTTCTACTAATAAATTTGGACAAGT	600
QY	601	AGTAACGGAAACGAAACGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAA	660
Db	601	AGCAACGGAAACGAAACGATGACACTGGATGCTTTGGGTACTCTTACGAGTAAATTTCT	660
QY	661	ACAAAGAACCAACATGTCTCGTCCAAATAATCTTCAACCACTGCCACAGCCCGTCCG	720
Db	661	ACGAAGAACCAACATGTCTCGTCTTAAACACCCCCACCGCCCAAGCCCGTCCG	720
QY	721	GAGATCAAACTACAAGCAACCACTGATGCCACCAAACTCAATACCAAGCAACCAAGC	780
Db	721	GAAATCAAAACCAAGCAACCACTGATGCCACTAGACTCAACACCAAAACCCCAAC	780
QY	781	AGTGTATGAGGACCTCGCAACATCTCGGCTCAGGCTCCGGGAAACGAGAACCCCAACA	840
Db	781	AGTGTATGATGAGGATCTCAACATCTCGGCTCGGGGTCTGGGGAAACAGGAATCTTATG	840
QY	841	ACTTCTGATGCGGTCAACGAAGGCTTTTCAACCAATGCCACCACTCCCTCACCAC	900
Db	841	ACTTCTGATGCGGTCACTAAGCAAGGCTTTTCAACCAATGCCACCACTCCCTCACC	900
QY	901	CAACCAAGCAACCCACAGCAAGGAGGAAACAAACCAATTTCCCAAGATGCTGTGATCT	960
Db	901	CAACCAAGCAACCCACAGCAAGGAGGAAACAAACCAATCTCCCAAGACGCTCAACTCT	960

QY	961	GAACCTAGACAAAAATAACACAACTGCACAAACCGCTCCATGCCCCCTCATAACTACCA	1020
DB	961	GAACCTGATTAACACCAACACAACTGCACAAACCGCCCGCTCCACAAACCA	1020
QY	1021	ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACAAAC	1080
DB	1021	ATCTCCACCAACACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACAAAC	1080
QY	1081	ACCACCAATGACACACAGAGACCAATCACTGAAAATGAGCAACACAGTGCCCTCG	1140
DB	1081	ACCAACCAATCCCAACACACAAAGCATGGTCACTGAAAATGAGCAACACAGTGCCCTCG	1140
QY	1141	ATAACAAACCTCTCCAAACCGGAAATCCCAACACAGCAAAAGAGCACAGCAGCAAAA	1200
DB	1141	AAAGCAACCTCTCCAAACCGGAAATCCCAACACAGCAAAAGAGCACCAACAATACAAA	1200
QY	1201	GGCCCCGACAAACCGGACCCCAACACAGCAAAATGAGCATTTCCAGAGTCCCTCC	1260
DB	1201	AGCCCCACCAATTTGGAACCAAAATACAAATGGAACATTTTCACTGCTCCCTCC	1260
QY	1261	CCAGCTCGAGTGCACAAACATCTTTGTATATTTTCAAGAAAGACGATCGATTTCTG	1320
DB	1261	CCCACTCGACTACAAACATCTTTATATATTTTCAAGAGAAACGAAATATCTCTG	1320
QY	1321	GAAGGTGATATATTTCCCGTTTTTATAGATGGGTTAATAAATCTGAAATTTGATCCA	1380
DB	1321	GAAGGTGATATATTTCCCGTTTTTATAGATGGGTTAATAAATCTGAAATTTGATCCA	1380
QY	1381	ATCCAAACACAGAAACATCTTTGATGAAATCTCCAGCTTTAATACTTCAACTAATG	1440
DB	1381	GTTCCAAATACAAAGACAACTTTTGTATGAAATCTTTCTAGTCTGCTGCTTCACT	1440
QY	1441	GAACAAACACTCCCGCAATATCAGTTTAACTTTCTCTTATTTTCTGTATAAAATGGA	1500
DB	1441	GATCAACATGATCTCTCAATATCAGTTTAACTTTAICTTATCTCTCATCAAGTGA	1500
QY	1501	GATCTGCTACTCTGCGGAAACCGAGAAATGATTTGTATGACAGAGTTTGAGGATTTG	1560
DB	1501	AACACTGCTACTCTGGAGAAATGAGAAATGATTTGTATGACAGAGCTAAGAAATTTG	1560
QY	1561	GTGAGAGGACGATTTGGCGGAGGCTTAGCTGGATACCACTTTTGGGCGCTGGAATC	1620
DB	1561	GTTCAGGAGGACGACTGGCAGCAGGCTCAGTTGGATACCACTTTTGGGCGCTGGAATC	1620
QY	1621	GAAGACTCTATCTGCGGTTTAAATCAAAATCAGAACTTTAGTTTGTAGGTTGAGG	1680
DB	1621	GAAGACTTTATACCGCAGGTTTAAATCAAAATCAGAACTTTGCTCGCAGGTTGAGG	1680
QY	1681	CGCTTAGCTAAATCAAACTGCTTAAATCTTGGAGCTCTTTGTTAAGGCTCAACACGAG	1740
DB	1681	CGTCTAGCCAAATCAAACTGCAAAATCTTTGGAACTCTTACTAAGGCTCAACACGAG	1740
QY	1741	AGGACATTTTCTTAATCAATAGCATGCAATGCACTTTTGTCTTACAGGCTGGGCGGA	1800
DB	1741	AGAACTTTTCTTAATCAATAGCATGCTATGCACTTTCTACTCAGAGTGGGAGGA	1800
QY	1801	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAAGCTTCTATAAAATATC	1860
DB	1801	ACATGCAAGGTGCTTGGACCTGATTTGTCATCGGAATAGAAAGCTTGTCCAGAAATAT	1860
QY	1861	TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGGTCTA	1920
DB	1861	TCAGAACAGATTGACCAATCAAGAAAGGACGAAACAAAGAGGAGGACTGGTGGGGTCTG	1920
QY	1921	GGTGGCAATGGTGGACATCTGACTGGGGTGTCTCACCATTTGGGCTCTCTGCTACTA	1980
DB	1921	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTTAACTTTGGGATTTTACTACTA	1980
QY	1981	TTATCTATAGTGTCTGATTTGCTCTGCTGTATCTCTGCTATCTTCACTAAATACATTT	2040
DB	1981	TTGTCCATAGTGTCTTGTATTTGCTCTATCTCTGATTTTGTCTGTAICTTTTACTAATATAT	2040

QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTTAAGGGTCACAAACGAGGAA 1740
DB 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTTAAGGGTCACAAACGAGGAA 1740
QY 1741 AGGACATTTTCCTTAATCAATAGGATGCAATTAAGTCTTTTGGCTTACGAGGTGGGGCGGA 1800
DB 1741 AGGACATTTTCCTTAATCAATAGGATGCAATTAAGTCTTTTGGCTTACGAGGTGGGGCGGA 1800
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
DB 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGGAGGAGAACTGGCTGGGGTCTA 1920
DB 1861 TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGGAGGAGAACTGGCTGGGGTCTA 1920
QY 1921 GGTGCAAAATGTTGACATCTGACTGGGGTGTTCACCAATTTGGGCATCTGCTACTA 1980
DB 1921 GGTGCAAAATGTTGACATCTGACTGGGGTGTTCACCAATTTGGGCATCTGCTACTA 1980
QY 1981 TTATCTATAGCTGTTCTGATTCCTCTGCTCTATCTGTCGATCTTCACTAAATACATT 2040
DB 1981 TTATCTATAGCTGTTCTGATTCCTCTGCTCTATCTGTCGATCTTCACTAAATACATT 2040
QY 2041 GGATGA 2046
DB 2041 GGATGA 2046
RESULT 13
AF005734
LOCUS
DEFINITION Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein
ACCESSION AF005734
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

KSLLDPPSNIRDYPKCTVHHIQONPHAGGIALHLWGAPFLYDRVASTTMYRKVF
TEGNIAAMI VNKTVHRMIFRQGGQYHMLNLTSTNKYWTSSNETORNDTGCFLQIYEY
NSTNDQCPFLSKPSPSLPTVPSIHTNTAKSGTMNPSSDEDELMISGSGBEQ
GPHTTLVAVTEKQSSITLSPSTSHQEQNSTPNSRSHVAVTNGTDTPTQPAFL
LNNNTPTVNTLYKNSLTPSPPTNTTNDTQRELAESEOTNAQLNLTLDPTBNPTT
GQDNTSLNIMTSDITSKHPTNSSPDSPPTPEPLYFRKRSI FWKEGDIPEPLDG
LINTEDFDPIPNTEITIDESPSTNTSEQHTPPNLSLTFSPFNKNGDTAVSGEN
ENDCADBLRWSQVEDLAAAGLSWIPFGPIEGLEYTAGLIKQNNLVCRLRRLANQ
AKSLELLRVTEERTFSLINRHAIDLTIRWGGTCKVLGPDCCIGIBDLKNSIQEQI
DKIRKQKEKETGMLGGKWTSDWGLVTNLGILLLSIAVLIALSCIRIFTKYIG"
1. .81
/gene="Gp"
82. .2043
/gene="Gp"
/product="glycoprotein"
1930. .2019
/gene="Gp"
/note="encodes transmembrane anchor peptide"
ORIGIN
Query Match 74.0%; Score 1513.2; DB 14; Length 2046;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTCCCTTATCAGTCTTATCTTAATTAATCAAGGACAAAAAATCTCCCC 60
DB 1 ATGAAGACCACATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAATCTCCCT 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120
DB 61 GTTTTAGAAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTCAAGTGTCTCCGGAAC 120
QY 121 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCAGAGACAGAGAGATGTTTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGGCATCCAAAGGATGGCTTTTCCAGGACAGGTGTACCTCCCAAGAT 240
DB 181 GATTCCCTTTGGAGGCATCTTAAACGATGGCTTTTCCAGGACAGGTGTCTCCCAAGAAC 240
QY 241 GTTCAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTACGATCC 300
DB 241 GTTCAGTATACGAGAGGAGAGAGCCAAACATGTTTACAATATAAGTGTACAGACCT 300
QY 301 TCTGAAAAATCCTTGCTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
DB 301 TCTGAAAAATCCTTGCTGTGATCTCCAGTATATATCCGATTTACCTTAATGTAA 360
QY 361 ACTATCCATCATATTCAGAGTCAAAACCCCTCATGACAGGGGATCGCCCTTCATTATGG 420
DB 361 ACTGTTCATCATATTTCAAGGTCAAAACCCCTCATGACAGGGGATTTGCCCTTCATTATGG 420
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCCAACAATGTACCGAGGAGAACTCTTC 480
DB 421 GGAGCATTTTCTGTATGATCGCGTTCCTCTACAACATGTACCGAGGAGAACTCTTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTCAATAAGACAGATGCAACAAATGATTTTCG 540
DB 481 ACTGAAGAAATATAGCAGCTATGATTTCAATAAGACAGATTCACAGATGATTTTCT 540
QY 541 CGGCAAGGACAAAGGTACCGTCAATATGAACTCTGATCTTACTATAAATATTTGGACAAGT 600
DB 541 AGGCAAGGACAAAGGTATTCGTACATGAACCTTGACCTCCACCAATAAATATTTGGACAAG 600
QY 601 AGTAACGGAACGCAAAACGAATGACATGATGTTTCCGGCTCTTCAAGATACAAATCT 660
DB 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTTCAAGATACAAATCT 660
QY 661 ACAAGGAAACCAAAACATGTGCTCCGTCCTCCAAATATACCTCCACCAAGCCCGCTCG 720
DB 661 ACAACATCAACATGACCTCCCTCATCTCTTAAACCTCCATCCCTGCCACAGTAAGTCCG 720
QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780

Db 721 AGCATCTCTACAAATCTCAAAATTAATTAATCTGTAATCTGGAACCTATGAACCCAAAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGCCCCACACA 840
Qy 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCATCAACAATGCCACCCACTCCCTCACA 900
Db 841 ACTCTTAATGTAGTCACCTGAACAGAAACAATGGTCAACAATATTGTCCACTCTCTTCACTA 900
Qy 901 CAACCAAGCAGCCACAGCAGAGGAGAAACAACAACCAACCTATCCCAAGATGCTGTGACT 960
Db 901 CATCCAGACCTCACACATGAGCAAAACAGTAGCAATCTTCCCGACATGCTGTAACT 960
Qy 961 GAACCTAGACAAATAACACAACTGCAACAACCGTCCATGCCCTCTCATPAACACTACCA 1020
Db 961 GAGCACAATGGAAACCGACCAACAACAACAGCAACGCTCTCAACAATAGTAATAACA 1020
Qy 1021 ATCTCTACTACACACACCTCCAAACACAACTTCAGCACTCTCTCTGCAACCTATCAAAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTAGCAACCTCAGTACTCTCTTCCCTCCCAACCGCAAC 1080
Qy 1081 ACCACCAATGACACACAGACACACATCACTGAAATGAGCAACAGTGGCCCTCG 1140
Db 1081 ATCACCAATATGATACACACAGTGAATAGCAGAAAGCGCAACAACCAATGCTCAGTTG 1140
Qy 1141 ATAAACAACCTGCTCCCAACGGGAAATCCCAACAGCAAGAGACCAAGCAGCAAAAAA 1200
Db 1141 AACACAACCTAGATCCAAAGAGAAATCCCAACAGCAAGACCAACCAAGCAACAC 1200
Qy 1201 GGGCCCGCACAAACGGCACCAACAGCAAAATGAGCAATTCACAGTCTCTCCCGCCAC 1260
Db 1201 AACATCATCATGACACATCATGATATAACAAGCAACACCCCAACAAATTTCTCCGGAT 1260
Qy 1261 CCCAGCTGACGACACACATCTGTATATTTCAGAGAAAGCGATCGATTTCTGGAA 1320
Db 1261 TCTAGTCGCAACCCCGCTCTTATATCTTTAGAAAGAACGAGCATTTCTCGAA 1320
Qy 1321 GAAGTGATATATTCCTGTTTTAGATGGGTAAATAACTGTAATTTGATTTGATCCA 1380
Db 1321 GAAGTGATATATTCCTGTTTTAGATGGGTAAATAACTGTAATTTGATTTGATCCA 1380
Qy 1381 ATCCCAACACAGAAACATCTTTGATGAATCTCCAGCTTTAATCTTCAACTAATGAG 1440
Db 1381 ATCCCAACACAGAAACATCTTTGATGAATCTCCAGCTTTAATCTTCAACTAATGAG 1440
Qy 1441 GAAACAACACTCCCGGCAATATCAGTTTAATCTTCTTATTTCTGATATAAATGGA 1500
Db 1441 GAAACAACACTCCCGGCAATATCAGTTTAATCTTCTTATTTCTGATATAAATGGA 1500
Qy 1501 GATCTGCTACTCTGGGGAACACAGAAATGATGTGATGAGTTCAGAGTTTGGAGT 1560
Db 1501 GATCTGCTACTCTGGGGAACACAGAAATGATGTGATGAGTTCAGAGTTTGGAGT 1560
Qy 1561 GTGAGGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTTGGCCCTGGAAATC 1620
Db 1561 GTGAGGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTTGGCCCTGGAAATC 1620
Qy 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Db 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Qy 1681 CGCTTAGCTTAATCAAACTGCTTAAATCTTTGGAGCTCTTTGTAAGGCTCACAAACGAGAA 1740
Db 1681 CGCTTAGCTTAATCAAACTGCTTAAATCTTTGGAGCTCTTTGTAAGGCTCACAAACGAGAA 1740
Qy 1741 AGGACATTTCTTAAATCAATAGGATCAATTTGCTTTACGAGGTGGGGCGGA 1800
Db 1741 AGGACATTTCTTAAATCAATAGGATCAATTTGCTTTACGAGGTGGGGCGGA 1800
Qy 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAATAATATC 1860

Db 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAATAATATC 1860
Qy 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGAGGAGAACTGGCTGGGGTCTA 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGAGGAGAACTGGCTGGGGTCTA 1920
Qy 1921 GGTGGCAATGGTGGACATCTGACTGGGGTCTTCCACCAATTTGGGCATCTCTGCTACTA 1980
Db 1921 GGTGGCAATGGTGGACATCTGACTGGGGTCTTCCACCAATTTGGGCATCTCTGCTACTA 1980
Qy 1981 TTATCTATAGTCTTCTGATTGCTCTGCTGTTATCTGCTGATCTTCACTAAATACATT 2040
Db 1981 TTATCTATAGTCTTCTGATTGCTCTGCTGTTATCTGCTGATCTTCACTAAATACATT 2040
Qy 2041 GGATGA 2046
Db 2041 GGATGA 2046

RESULT 14
AX717719
LOCUS AX717719 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 7 from Patent WO02079239.
ACCESSION AX717719
VERSION . AX717719.1 GI:29890728
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
REFERENCE 1 Chimeric filovirus glycoprotein
AUTHORS Patent: WO 02079239-A 7 10-OCT-2002;
TITLE U.S. Army Medical Research Institute of Infectious Diseases (US)
JOURNAL Location/Qualifiers
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Chimeric molecule between Marburg virus strain
Raven Glycoprotein 1 and Marburg virus strain Musoke
Glycoprotein 2"

ORIGIN
Query Match 64.5%; Score 1319.6; DB 6; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;
Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCC 60
Db 1 ATGAAGACCAATATTTCTGATTAGTCTCATTTTATCCAAAGTATAAAACTCTCCCT 60
Qy 61 ATTTTAGATAGTCTAGTAAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAAATTTGCTAGTAACAGCCCAACCTCAAGATGTAGATTCTGCTGCTCCGGAAC 120
Qy 121 CTCAGAGACAGAAAGACGTCCTCATCTGATGGGATTTCACACTGAGTGGGCAAAAAGTGTCT 180
Db 121 CTCAGAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTGTCT 180
Qy 181 GATTCCCTTTGGAGGCATCCAAAGCATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAAAT 240
Db 181 GATTCCCTTTGGAGGCATCTAAACGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAAAT 240
Qy 241 GTTGAGTACAGAGGGGGAGAGCCAAACATGCTACATATAGTGTACCGATCCC 300
Db 241 GTTGAGTATACGAGAGGAGAGAGCCAAACATGTTTCAATATAAGTGTACAGACCT 300
Qy 301 TCTGAAAAATCTTGTGCTTGTAGATCTCTCAACACTCCGCTACTATCCGAAATGCAAA 360
Db 301 TCTGAAAAATCTTGTGCTGCTGATCTCTCCAGTAAATATCCCGGATTACCTTAATGTAAA 360
Qy 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCTTATATGG 420

Db	361	ACTGTTTCATCATATTAATCAAGGTCAAAACCCCTCATGCGACAGGGGATGCCCTCCATTTGTGTGG	420
Qy	421	GGAGCAATTTTTCGTGTATGATCGCATTTGCGTCCACAACAATGATPACGAGCGAAAGTCTTTC	480
Db	421	GGGGCATTTTTCGTGTATGATCGCGTTGCGTCTACAAACAATGATPACGAGCGCAAGTCTTTC	480
Qy	481	ACTGAAGGGGAACATAGCAGCTATGATTTGTCAATPAGACAGTGCACAAAATGATTTTCTCG	540
Db	481	ACTGAAGGAAATATAGCAGCTATGATTTGTAAATAGACAGTTCACAGAATGATTTTCTCT	540
Qy	541	CGGCAAGGACAAGGTTACCGTCAATATGAATCTGACATTCTACTTAATAAATATTGGACAAGT	600
Db	541	AGGCNAGGACAAGGTTATCGTGCATGAACTTGACCTCCACCAATAAATATTGGACAAGC	600
Qy	601	AGTAAAGGAAACGAAAGCAATGACACTGGATGTTTGGGGCTCTTCAAGAATAACAATTTCT	660
Db	601	AGCAATGAACGCGAGAAATGATACGGATGTTTGGCATCTCTCCAGNATACAACCTCC	660
Qy	661	ACAAAGAACCAAAACATGTGTCCGTCCMAAATACTCCACACTGCGCCACAGCCCGTCGG	720
Db	661	ACAAACAATCAAAACATGCGCTCCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG	720
Qy	721	GAGATCAAACCTCAAGCACCCCAACTGATGCGCACCAAACTCAATACCAGGACCCCAAGC	780
Db	721	AGCATTCACCTCTCAAAATACTCAAAATTAATACTGCTAAATCTGGAACTATGAACCCCAAGT	780
Qy	781	AGTGATGATCAGGACCTCGCAACATCGGGCTCAGGGTCGGGAGAACGAGAAACCCCAACA	840
Db	781	AGCGCAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGGCCCCACACA	840
Qy	841	ACTTCTGATCGGGTCACAAAGCAGGGCTTTCATCAACAATGCGCACCCACTCCCTCACCA	900
Db	841	ACTCTTAATGTAGTCTACTGAAAGAAACAATCGTCAACAATATTGTCCACTCTCTTCACTA	900
Qy	901	CAACCAAGCAGCCACAGCAGGAGGAAACAAACACAAACCATTTCCCAAGATGCTGTGACT	960
Db	901	CATCCAAGCACCTCACAAACATGAGCAAAAACAGTACGAATCCTTTCCGACATGCTGTAACT	960
Qy	961	GAACTAGACAAAATTAACAACTGCGACACCGTCCATGCGCCCTCATACACTACCACA	1020
Db	961	GAGCACAATGGAAACCGACCCCAACAACAACACAGCAACGCTCTCTCAACAATATAATACA	1020
Qy	1021	ATCTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACCATTTACAAAAC	1080
Db	1021	ACTCCCACTTATAACACTCTCAAGTACAACTCAGTACTCTTTCCTCCCAACCCGCAAC	1080
Qy	1081	ACCACCAATGACAAACACAGAGCACAATCAGCTGAAAATGAGCAAAACGAGTGCCTCTCG	1140
Db	1081	ATCACCATAATGATACACAACGCTGAACTAGCAGAAAGCGCAACAAACCAATGCTCAGTTG	1140
Qy	1141	ATAACACCTCGCTCCAAACGGGAAATCCCAACACAGCAAGAGAGCACCAGCAGCAAAAA	1200
Db	1141	AACACAACCTTAGATCCAAAGAAAAATCCCAACACAGGACAGACACCAAGCAGCAACCC	1200
Qy	1201	GGCCCCGCCACAACGGCACCAACACAGCAAAATGAGATTTTCACAGTCTCTCCCCCACC	1260
Db	1201	AACATCATATGACGATCAGATATTAAGAAGCAACACCCCAACAAATTTCTTCGGAT	1260
Qy	1261	CCCAGCTCGACTGCACAACATCTTTGTATATTTTCAGAAAGAAAGCGATCGATTTTCTGGAAA	1320
Db	1261	TCTAGTCCGACAAACCCGCCCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG	1320
Qy	1321	GAAGGTGATATATTCGGTTTTTATAGATGGGTTAATAAATACTGAAATTTGATTTGATCCA	1380
Db	1321	GNAGGCGACATGTTTCCCTTTTCTGGATGGGTTAATAAATGCTCCAAATGATTTTGACCCA	1380
Qy	1381	ATCCCAACACAGAAAACAACTTTTGTAGTAATCTCCAGCTTTAATACTCTCAACTAATAG	1440
Db	1381	GTTTCCAATATACAAAAACAATCTTTGTAGTAATCTCTAGTTCTGGTGTCTCGGTGAGGAA	1440
Qy	1441	GAACAAACACACTCCCCCGAATATCAGTTTAACTTTTCTTTATTTTCTGTATAAAAAATGGA	1500

1441	DB	GATCAACATGCCCTCCCCCAATATTAGTTTAACTTTATCTTATTTCTTAATAAATGAG	1501
1501	QY	GATACTGCCCTACTCTGGGGAACGAGAATGATTTGATGTCAGAGTTGAGGATTTGGAGT	1560
1501	DB	AACACTGCCCTACTCTGGAGAAAATGAGAATGATTTGTGATGTCAGAGTTAAGAAATTTGGAGC	1560
1561	QY	GTGAGGAGACAGATTTGGCGGCAGGGCTTAGCTGTGATACCAATTTTTTGGCCCTGGAAATC	1620
1561	DB	GTTTCAGGAGGATGACCTGGCGGCAGGGCTCAGTTGTGATACCGTTTTTTTGGCCCTCGGAATT	1620
1621	QY	GAAGGACTCTATACTGCGCGTTTTAAATCAAAAAATCAGAACAAATTTAGTTTGTAGGTTTGAGG	1680
1621	DB	GAAGGACTTTTACACTGCTGTTTTTAAATTAATAATCAAAACAAATTTGGTCTGCGAGTTTGAGG	1680
1681	QY	CGTTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTTGTTAAGGTCACAACCGAGGAA	1740
1681	DB	CGTCTAGCCAAATCAAACTGCGAAAATCCTTTGGAACTCTTATTGAGAGTTCACAACCTGAGGAA	1740
1741	QY	AGGACATTTTCCTTAATCAATAGGCATGCAATTTGACTTTTTTGTCTACGAGGTGGGGCGGA	1800
1741	DB	AGAACATTTTCCTTAATCAATAGACATGCTATTGACTTTCTACTCAACAGATGGGGAGGA	1800
1801	QY	ACATGCAAGGTGCTTAGGACCTGATTTGTCATAGGAATAGAAGATCTATCTAAAAATATC	1860
1801	DB	ACATGCAAAAGTGCTTGAGCCTGATTTGTCATCGGGATAGAAGACTTGTCCAAAAATATT	1860
1861	QY	TCAGAACAAATTCGACAAAATCAGAAAGGATGAAACAAAGAGGAGAACTGCTGGGTCTA	1920
1861	DB	TCAGAGCAAATTTGACCAAATTTAAAAAGGACGAAACAAAAAGAGGGGACTGGTTGGGGTCTG	1920
1921	QY	GGTGGCAAATGGTGGACATCTGACTGGGGTGTCTCACCAATTTGGGCATCCTGCTACTA	1980
1921	DB	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGTCTACTA	1980
1981	QY	TTATCTATAGCTGTTCTGATTTGCTCTGCTCTGATCTGTCTGATCTTCTCACTAAATACATT	2040
1981	DB	TTATCCATAGCTGTCCTGATTTGCTCTATCTCTGATTTGTCTGATCTTTTACTTAATATATC	2040
2041	QY	GGATGA 2046	
2041	DB	GGATAA 2046	

RESULT 15			
AX717715			
LOCUS	AX717715	1841 bp	DNA
DEFINITION	Sequence 3 from Patent WO02079239.	linear	PAT 15-APR-2003

AX717715.1 GI:29890726

KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
other sequences; artificial sequences.

REFERENCE
AUTHORS
1 Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.

TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 3 10-OCT-2002;

FEATURES

U.S. Army Medical Research Institute of Infectious Diseases (US)
Location/Qualifiers

source

```

1: 1011
source
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

```

[/notes="chimeric molecule between Marburg virus strain
Musoke Glycoprotein 1 and Ebola virus Zaire Mayinga strain
Glycoprotein 2"](#)

ORIGIN

Query Match 63.9%; Score 1308; DB 6; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308: Conservative 0; Mismatches 0; Indels 0

1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAAGGACAAAATCTCCCC 60


```
Db 1 ATGAAGACACATGTTTCCTTATCAGTCTTAATCTTAATCAAGGGCAAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
QY 121 CTCCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGCTCT 180
Db 121 CTCCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGCTCT 180
QY 181 GATTCCTCTTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCTCTTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
QY 241 GTTAGATACAGAGGGGAGAGGAGCAAAAATCATGCTACATATAAGTGTAAACGATCCC 300
Db 241 GTTAGATACAGAGGGGAGAGGAGCAAAAATCATGCTACATATAAGTGTAAACGATCCC 300
QY 301 TCTGGAAAATCCTTCTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAAAATCCTTCTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTATGG 420
QY 421 GGAGCATTTTTCTGTATGATCGCATGTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCATGTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
QY 481 ACTGAAGGACATAGCAGCTATGATGCTCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db 481 ACTGAAGGACATAGCAGCTATGATGCTCAATAGACAGTGCACAAAATGATTTTCTCG 540
QY 541 CGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTATAAATATTGGACAAGT 600
Db 541 CGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTATAAATATTGGACAAGT 600
QY 601 AGTAAACGGAACGCAAAAGTGAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT 660
Db 601 AGTAAACGGAACGCAAAAGTGAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT 660
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACACTGCCACAGCCCGTCCG 720
Db 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACACTGCCACAGCCCGTCCG 720
QY 721 GAGATCAAACTCACAAAGCACCCCACTGATGCCACCAAACTCAATACCAGGACCCCAAGC 780
Db 721 GAGATCAAACTCACAAAGCACCCCACTGATGCCACCAAACTCAATACCAGGACCCCAAGC 780
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
QY 841 ACTTCTGATGGGTCAACAAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA 900
Db 841 ACTTCTGATGGGTCAACAAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA 900
QY 901 CAACCAAGCAGCCACAGCAGGAGGAAACACACAAACCATTCCTCCAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCACAGCAGGAGGAAACACACAAACCATTCCTCCAGATGCTGTGACT 960
QY 961 GAACTAGACAAAATAACAACTGCACAAACCGTCCATGCCCTCTCATACAACTACCACA 1020
Db 961 GAACTAGACAAAATAACAACTGCACAAACCGTCCATGCCCTCTCATACAACTACCACA 1020
QY 1021 ATCTCTACTAACAAACCTTCAAAACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
Db 1021 ATCTCTACTAACAAACCTTCAAAACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
QY 1081 ACCACCAATGACACAGAGGACATCACTGAAAATGAGCAACCCAGTGCCTCCCTCG 1140
```

```
Db 1081 ACCACCAATGACACAGAGGACATCACTGAAAATGAGCAACCCAGTGCCTCCCTCG 1140
QY 1141 ATAAACAACCTTGCCTTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGAGCAAAAAA 1200
Db 1141 ATAAACAACCTTGCCTTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGAGCAAAAAA 1200
QY 1201 GGCCCGGCCACAACGGCACCAACACAGCAAAATGAGCAATTTTCAACAGTCTCTCCCTCCACC 1260
Db 1201 GGCCCGGCCACAACGGCACCAACACAGCAAAATGAGCAATTTTCAACAGTCTCTCCCTCCACC 1260
QY 1261 CCCAGCTCGACTGCACACATCTTTGTATATTTTCAGAGAAAGCGATCG 1308
Db 1261 CCCAGCTCGACTGCACACATCTTTGTATATTTTCAGAGAAAGCGATCG 1308
```

Search completed: August 1, 2005, 02:43:48
Job time : 8800.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1071.87 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-5

Perfect score: 2046

Sequence: 1 atgaagaccacatgttcct.....tcactaaatacatggatga 2046

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	10	ABT13452 DNA encod
2	1841.2	90.0	2103	8	ACC57764 Marburg v
3	1841.2	90.0	2164	4	AAD04042 Marburg v
4	1841.2	90.0	7778	10	ACC71550 VRC6701 (
5	1841.2	90.0	11460	3	Az87211 VEE repli
6	1825.2	89.2	2100	3	Az51039 Marburg v
7	1813.6	88.6	2051	10	ABT13454 DNA encod
8	1767.6	86.4	7005	10	ACC71551 VRC6702 (
9	1767.6	86.4	8256	10	ACC71552 VRC6710 (
10	1692.4	82.7	19112	13	ADM48345 Marburg v
11	1518	74.2	2046	10	ABT13456 DNA encod
12	1510	73.8	2247	4	AAD04043 Marburg v
13	1313.6	64.5	2046	10	ABT13453 DNA encod
14	1308	63.9	1841	10	ABT13451 DNA encod
15	995.6	48.7	6902	10	ACC71562 VRC6703 (p
16	546.6	26.7	2252	10	ABT13450 DNA encod
17	145	7.1	6940	10	ACC71535 VRC6202 (
18	145	7.1	7073	2	AAX59391 Plasmid p
19	145	7.1	7082	10	ACC71533 VRC6200 (
20	145	7.1	7087	10	ACC71534 VRC6201 (

21	145	7.1	8221	10	ACC71547	Acc71547 VRC6602 (
22	145	7.1	8338	10	ACC71546	Acc71546 VRC6601 (
23	145	7.1	10783	10	ACC71545	Acc71545 VRC6600 (
24	143.4	7.0	7073	2	AAX89796	Aax89796 DNA of pV
25	140.4	6.9	7001	2	AAX89795	Aax89795 DNA of pV
26	138.8	6.8	6885	10	ACC71538	Acc71538 VRC6302 (
27	138.8	6.8	6889	10	ACC71539	Acc71539 VRC6303 (
28	138.8	6.8	7002	10	ACC71536	Acc71536 VRC6300 (
29	138.8	6.8	7003	2	AAX59390	Aax59390 Plasmid p
30	138.8	6.8	7023	10	ACC71541	Acc71541 VRC6351 (
31	138.8	6.8	7036	10	ACC71537	Acc71537 VRC6301 (
32	138.8	6.8	8146	10	ACC71540	Acc71540 VRC6310 (
33	131.2	6.4	2039	10	ABT13455	Abt13455 DNA encod
34	131.2	6.4	2172	4	AAD04041	Aad04041 Ebola vir
35	131.2	6.4	2224	3	AAX51038	Aax51038 Ebola vir
36	131.2	6.4	2298	3	AAX87189	Aax87189 Ebola vir
37	131.2	6.4	2298	4	AAF76953	Aaf76953 Ebola vir
38	131.2	6.4	2298	12	ADL27460	Adl27460 Nucleotid
39	131.2	6.4	6467	10	ACC71530	Acc71530 VRC6052 (
40	131.2	6.4	6561	10	ACC71524	Acc71524 VRC6003 (
41	131.2	6.4	6624	10	ACC71523	Acc71523 VRC6002 (
42	131.2	6.4	6914	10	ACC71529	Acc71529 VRC6008 (
43	131.2	6.4	7154	10	ACC71521	Acc71521 VRC6000 (
44	131.2	6.4	7188	10	ACC71522	Acc71522 VRC6001 (
45	131.2	6.4	7272	2	AAX59393	Aax59393 Plasmid p

ALIGNMENTS

RESULT 1

ABT13452
ID ABT13452 standard; DNA; 2046 BP.

XX ABT13452;

XX 30-JAN-2003 (first entry)

XX DNA encoding a chimeric filovirus protein MUS-GP1/RVN-GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

XX immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

XX Chimeric.

XX WO200279239-A2.

XX 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.

XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX Grogan CC, Hevey MC, Schmaljohn AL;

XX WPI; 2003-040651/03.

XX P-FSDB; ABJ18474.

XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2, useful for inducing an immune response against infection of different filoviruses, specifically against both Ebola and Marburg viruses.

XX Claim 16; Page 74-76; 94pp; English.

XX The invention relates to a chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2, where GP1 is from a filovirus different than that of GP2. The chimeric filovirus GP protein is useful for inducing an immune response against infection of different filoviruses, specifically against both Ebola and Marburg viruses by being used as a vaccine. This polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention	
XX		
SQ	Sequence 2046 BP; 659 A; 517 C; 402 G; 468 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2046; DB 10; Length 2046;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAAGACCACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	60
DB	1 ATGAAGACCACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	60
QY	61 ATTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT	120
DB	61 ATTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT	120
QY	121 CTCGAGACAGACAGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGCTT	180
DB	121 CTCGAGACAGACAGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGCTT	180
QY	181 GATTCCTCTTGGAGCATCCAAAGCATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
DB	181 GATTCCTCTTGGAGCATCCAAAGCATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
QY	241 GTTGAGTACACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	300
DB	241 GTTGAGTACACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	300
QY	301 TCTGGAATTCCTTGTGTAGATCTCTTACCAATCCCTGAGCTATCCGAAATGCAAA	360
DB	301 TCTGGAATTCCTTGTGTAGATCTCTTACCAATCCCTGAGCTATCCGAAATGCAAA	360
QY	361 ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTATGG	420
DB	361 ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTATGG	420
QY	421 GGAGCATTTTTCTGTATGATCGCATGCTCTCCACAAATGATACCGAGGCAAAAGTCTTC	480
DB	421 GGAGCATTTTTCTGTATGATCGCATGCTCTCCACAAATGATACCGAGGCAAAAGTCTTC	480
QY	481 ACTGAAGGGAACATAGCAGCTATGATGTCAATAGACAGTGCACAAATGATTTTCTCG	540
DB	481 ACTGAAGGGAACATAGCAGCTATGATGTCAATAGACAGTGCACAAATGATTTTCTCG	540
QY	541 CGGCAAGGAACAGGGTACCGTCTATGATCTGACTTCTACTAATAAATATTGGCAAGT	600
DB	541 CGGCAAGGAACAGGGTACCGTCTATGATCTGACTTCTACTAATAAATATTGGCAAGT	600
QY	601 AGTAACGGGAACGAAACGAATGACACTGCTGCTTTCGGGCTCTTCAAGAATACAAATCT	660
DB	601 AGTAACGGGAACGAAACGAATGACACTGCTGCTTTCGGGCTCTTCAAGAATACAAATCT	660
QY	661 ACAAGAACCAACATGCTCGTCCAAATACCTCAGCACTGCTCCACAGCCCGTCCG	720
DB	661 ACAAGAACCAACATGCTCGTCCAAATACCTCAGCACTGCTCCACAGCCCGTCCG	720
QY	721 GAGATCAAACTCACAGACACCCCACTGATGCCACCAAACTCAATACACGACCCCAAGC	780
DB	721 GAGATCAAACTCACAGACACCCCACTGATGCCACCAAACTCAATACACGACCCCAAGC	780
QY	781 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	840
DB	781 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	840
QY	841 ACTTCTGATGGGTACCAAGCAAGGGCTTTTCAATCAAAATGCTCCCTCCCTCACCA	900
DB	841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCAATCAAAATGCTCCCTCCCTCACCA	900
QY	901 CAACCAAGCACGCCACAGCAAGGGAACACAAACCAATTCCTCAAGATGCTGTGACT	960
DB	901 CAACCAAGCACGCCACAGCAAGGGAACACAAACCAATTCCTCAAGATGCTGTGACT	960
QY	961 GAACTAGACAAAAATAACAACTGACAAACCGTCCATGCCCCCTCATAACTACCACA	1020
DB	961 GAACTAGACAAAAATAACAACTGCAACCGTCCATGCCCCCTCATAACTACCACA	1020
QY	1021 ATCTCTACTTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTATCAAAAC	1080
DB	1021 ATCTCTACTTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTATCAAAAC	1080
QY	1081 ACCACCAATGATCAACACACAGAGACCAATCACTGAAAAATGAGCAAAACAGTGCCTCCG	1140
DB	1081 ACCACCAATGATCAACACACAGAGACCAATCACTGAAAAATGAGCAAAACAGTGCCTCCG	1140
QY	1141 ATACCAACCTTGCCTCCAAACGGGAAATCCCAACAGCAAGCAAGAGACCAAGCAAGAAAA	1200
DB	1141 ATACCAACCTTGCCTCCAAACGGGAAATCCCAACAGCAAGCAAGAGACCAAGCAAGAAAA	1200
QY	1201 GGGCCGCGCACAAAGGGACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCCCACC	1260
DB	1201 GGGCCGCGCACAAAGGGACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCCCACC	1260
QY	1261 CCCAGCTCGACTGCACACAACTCTTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA	1320
DB	1261 CCCAGCTCGACTGCACACAACTCTTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA	1320
QY	1321 GAAAGTGATATATTTCCCGTTTATAGTGGGTTAATAATACTGAAATTTGATTTGATCCA	1380
DB	1321 GAAAGTGATATATTTCCCGTTTATAGTGGGTTAATAATACTGAAATTTGATTTGATCCA	1380
QY	1381 ATCCCAACACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACCTTCACTAATGAG	1440
DB	1381 ATCCCAACACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACCTTCACTAATGAG	1440
QY	1441 GAAACAAACACTCCCGGAAATATCAGATTTAACTTTCTCTATTTTCTGATAAAATGGA	1500
DB	1441 GAAACAAACACTCCCGGAAATATCAGATTTAACTTTCTCTATTTTCTGATAAAATGGA	1500
QY	1501 GATATGCTTACTCTGCGGAAAAACGAGAAATGATGATGACAGAGTTGGAGT	1560
DB	1501 GATATGCTTACTCTGCGGAAAAACGAGAAATGATGATGACAGAGTTGGAGT	1560
QY	1561 GTGAGGAGGAGGATTTGGCGGAGGCTTACGTGATACCAATTTTGGCCCTCGAATC	1620
DB	1561 GTGAGGAGGAGGATTTGGCGGAGGCTTACGTGATACCAATTTTGGCCCTCGAATC	1620
QY	1621 GAAAGACTCTATCTGCTGCGGTTTAAATCAAAATCAGAAACAAATTTAGTTTGTAGTTGAGG	1680
DB	1621 GAAAGACTCTATCTGCTGCGGTTTAAATCAAAATCAGAAACAAATTTAGTTTGTAGTTGAGG	1680
QY	1681 CGTTAGCTAATCAAACTGCTGCTTAAATCTTTGGAGCTCTTTGTTAAGGCTCACAACCGAGAA	1740
DB	1681 CGTTAGCTAATCAAACTGCTGCTTAAATCTTTGGAGCTCTTTGTTAAGGCTCACAACCGAGAA	1740
QY	1741 AGGACATTTTCTTAACTAATAGCATGCAATTTGCTTTTACGAGGTGGGGCGGA	1800
DB	1741 AGGACATTTTCTTAACTAATAGCATGCAATTTGCTTTTACGAGGTGGGGCGGA	1800
QY	1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAATAATATC	1860
DB	1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAATAATATC	1860
QY	1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA	1920
DB	1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA	1920
QY	1921 GGTGGCAATGGTGACATCTGCTGGGGTCTCTCACCATAATTTGGGATCTCTGCTACTA	1980
DB	1921 GGTGGCAATGGTGACATCTGCTGGGGTCTCTCACCATAATTTGGGATCTCTGCTACTA	1980
QY	1981 TTATCTATAGTGTCTGATTTGCTCTGCTGATCTCTGCTGATCTCTTCACTAATAATATC	2040
DB	1981 TTATCTATAGTGTCTGATTTGCTCTGCTGATCTCTGCTGATCTCTTCACTAATAATATC	2040
QY	2041 GGATGA 2046	

Db 2041 GGATCA 2046

RESULT 2
ACC57764
ID ACC57764 standard; cDNA; 2103 BP.
XX
AC ACC57764;
DT 28-JUL-2003 (first entry)
XX
DE Marburg virus envelope glycoprotein coding sequence.
XX
KW Lentivirus; pseudotyped virus; gene therapy; vector; gene; ss.
XX
OS Marburg virus.
FH Key Location/Qualifiers
FT CDS 10..2055
FT /*tag= a
FT /product= "Envelope glycoprotein"
XX
PN WO2003035849-A2.
XX
PD 01-MAY-2003.
XX
PF 28-OCT-2002; 2002WO-US034545.
XX
PR 26-OCT-2001; 2001US-0353221P.
PR 26-OCT-2001; 2001US-0356436P.
XX
PA (MCCR/) MCCRAY P B.
PA (SAND/) SANDERS D A.
PA (DAVI/) DAVIDSON B L.
XX
PI Mcclay PB, Sanders DA, Davidson BL;
XX
DR WPI; 2003-421416/39.
DR P-PSDB; ABR42242.
XX
PT New pseudotyped lentivirus comprising a lentiviral capsid, a lipid bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful for eliciting an immune response against feline immunodeficiency virus infection.
PT
PS Disclosure; Page 36-39; 4lpp; English.
XX
CC The present sequence is the coding sequence for the envelope glycoprotein of Marburg virus. The invention provides methods for gene transfer to cells using glycoprotein-pseudotyped lentiviruses. The glycoprotein is preferably Marburg glycoprotein. The Marburg viral glycoprotein preferably has a mutation in the C-terminal portion (see ABR42243-47) that results in a higher titre production of the pseudotyped virus. The pseudotyped lentivirus is used to introduce nucleic acid sequences encoding a desired protein into a hepatocyte, brain glial or airway epithelial cell, in vitro or in vivo. The desired protein is preferably cystic fibrosis transmembrane conductance regulator, or is the low density lipoprotein receptor, alpha1-antitrypsin, ornithine transcarbamylase, Factor VIII or a high affinity glutamate receptor (all claimed)
XX
SQ Sequence 2103 BP; 671 A; 533 C; 408 G; 491 T; 0 U; 0 Other;

Query Match 90.0%; Score 1841.2; DB 8; Length 2103;
Best Local Similarity 93.7%; Pred No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 60
DB 10 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 69
QY 61 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAC 120

Db 70 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAC 129
QY 121 CTCAGAAAGACAGAAAGAGCTCCATCTCATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
DB 130 CTCAGAAAGACAGAAAGAGCTCCATCTCATGGGATTCACATGAGTGGGCAAAAAGTTGCT 189
QY 181 GATTCCCTTTGGAGGATCCNAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
DB 190 GATTCCCTTTGGAGGATCCNAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 249
QY 241 GTTGAGTACACAGAGGGGAGAGCCAAAACATGCTACATAATAAGTGTAAACGGATCCC 300
DB 250 GTTGAGTACACAGAGGGGAGAGCCAAAACATGCTACATAATAAGTGTAAACGGATCCC 309
QY 301 TCTGAAAAATCCTTTGCTTTAGATCTCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
DB 310 TCTGAAAAATCCTTTGCTTTAGATCTCTCTACCAACATCCGTGACTATCCGAAATGCAAA 369
QY 361 ACTATCCATCATATTCAAGGTCAAAAACCTCTNTGACAGGGGATCGCCCTTCATTTATGG 420
DB 370 ACTATCCATCATATTCAAGGTCAAAAACCTCTNTGACAGGGGATCGCCCTTCATTTATGG 429
QY 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGGCAAAAGTCTTC 480
DB 430 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGGCAAAAGTCTTC 489
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAAATGATTTTCG 540
DB 490 ACTGAAGGGAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAAATGATTTTCG 549
QY 541 CGGCAAGGACNAGGGTACCGTCAATGATCTGACTTCTACTAATTAATATATGGACAGT 600
DB 550 CGGCAAGGACNAGGGTACCGTCAATGATCTGACTTCTACTAATTAATATATGGACAGT 609
QY 601 AGTAACGGAACGCAAAATGACATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATCT 660
DB 610 AGTAACGGAACGCAAAATGACATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATCT 669
QY 661 ACAAGAAACCAACATGTGCTCCGTCCAAAATACCTCCACACTGCCCCACAGCCGCTCG 720
DB 670 ACAAGAAACCAACATGTGCTCCGTCCAAAATACCTCCACACTGCCCCACAGCCGCTCG 729
QY 721 GAGATCAAACTCACAGCACCCCAACTGATGTCGACCAAACTCAATACACGAGCCCAAGC 780
DB 730 GAGATCAAACTCACAGCACCCCAACTGATGTCGACCAAACTCAATACACGAGCCCAAGC 789
QY 781 AGTCATGATGAGGACCTCGCAACATCCGGCTCAGGTCGAGAACGAGAACCCACACA 840
DB 790 AGTCATGATGAGGACCTCGCAACATCCGGCTCAGGTCGAGAACGAGAACCCACACA 849
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 900
DB 850 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 909
QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAAACAACAACCAATTCCTCAAGATGCTGTGACT 960
DB 910 CAACCAAGCAGCCACAGCAAGGAGGAAAACAACAACCAATTCCTCAAGATGCTGTGACT 969
QY 961 GAACCTAGCAAAAAATAACACAACCTGCAACCGCTCATGCCCCCTCATATAACACTACCACA 1020
DB 970 GAACCTAGCAAAAAATAACACAACCTGCAACCGTCCATGCCCCCTCATATAACACTACCACA 1029
QY 1021 ATCTCTACTAAACAACACCTCCAAAACAACACTTTCAGCACTCTCTCTGCAACCAATTAACAAC 1080
DB 1030 ATCTCTACTAAACAACACCTCCAAAACAACACTTTCAGCACTCTCTCTGCAACCAATTAACAAC 1089
QY 1081 ACCACCAATGACAACAACAGAGACCAATCACTGAAATAGAGCAAAACAGTGCCTCCCTCG 1140
DB 1090 ACCACCAATGACAACAACAGAGACCAATCACTGAAATAGAGCAAAACAGTGCCTCCCTCG 1149
QY 1141 ATACACACCTGCTCCCAACGGGAAATCCCAACAGCAGCAAGAGCAGCAGGACGCAAAAA 1200
DB 1150 ATACACACCTGCTCCCAACGGGAAATCCCAACAGCAGCAAGAGCAGCAGGACGCAAAAA 1209

QY 361 ACTATCCATCATATTTCAAGGTCAAAAACCTCATGCAAGGGGATCGCCCTTCATTATGG 420
Db 479 ACTATCCATCATATTTCAAGGTCAAAAACCTCATGCAAGGGGATCGCCCTTCATTATGG 538
QY 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
Db 539 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 598
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 599 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTCTCG 658
QY 541 CGGCAGGACAGGATACCGTCAATAGTATGATTTGTCAATAGACAGTGCACAAAATGATTTCTCG 600
Db 659 CGGCAGGACAGGATACCGTCAATAGTATGATTTGTCAATAGACAGTGCACAAAATGATTTCTCG 718
QY 601 AGTAACGGAACGCAAAACCAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATTTCT 660
Db 719 AGTAACGGAACGCAAAACCAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATTTCT 778
QY 661 ACAAGAAACCAAAATGTCCTCGTCCAAAATACCTCCACACACTGCCACAGCCCGTCCG 720
Db 779 ACAAGAAACCAAAATGTCCTCGTCCAAAATACCTCCACACACTGCCACAGCCCGTCCG 838
QY 721 GAGATCAAACTCAGACACCCCAACTGATGCGACCAAACTCAATACACGGAACCCAGC 780
Db 839 GAGATCAAACTCAGACACCCCAACTGATGCGACCAAACTCAATACACGGAACCCAGC 898
QY 781 AGTGATGATGAGGACCTCGCAACATCGCGCTCAGGGTCGGAGAACGAGAACCCACACA 840
Db 899 AGTGATGATGAGGACCTCGCAACATCGCGCTCAGGGTCGGAGAACGAGAACCCACACA 958
QY 841 ACTTCTGATGGGTCAACAAGCAAGGGCTTTTCATCAACAATGSCACCACTCCCTCACCA 900
Db 959 ACTTCTGATGGGTCAACAAGCAAGGGCTTTTCATCAACAATGSCACCACTCCCTCACCA 1018
QY 901 CAACAAGCAGCCACAGCAAGGAGGAAACAAACAACCAATTCGCAAGATGCTGTGACT 960
Db 1019 CAACAAGCAGCCACAGCAAGGAGGAAACAAACAACCAATTCGCAAGATGCTGTGACT 1078
QY 961 GAACTAGACAAATAAACAACAACATGCAACCGTCCATGCGCCCTCATACACTACCCACA 1020
Db 1079 GAACTAGACAAATAAACAACAACATGCAACCGTCCATGCGCCCTCATACACTACCCACA 1138
QY 1021 ATCTCTACTAACAACACTCTCAAAACACAACTTCAGCACTCTCTCTGCAACCAATTACAAAAC 1080
Db 1139 ATCTCTACTAACAACACTCTCAAAACACAACTTCAGCACTCTCTCTGCAACCAATTACAAAAC 1198
QY 1081 ACCACCAATGACACACACAGACACAATCACTGAAAATGAGCAAAACAGTGCCCTCTCG 1140
Db 1199 ACCACCAATGACACACACAGACACAATCACTGAAAATGAGCAAAACAGTGCCCTCTCG 1258
QY 1141 ATAACAACCTGCTCCCAACGGAATCCCAACAGCAGCAAGAGCAACAGCAGCAAAAAA 1200
Db 1259 ATAACAACCTGCTCCCAACGGAATCCCAACAGCAGCAAGAGCAACAGCAGCAAAAAA 1318
QY 1201 GSCCCCGGCAACAACGCGACCAACAGCAATGAGCAATTCACAGTCTCCCTCCCAAC 1260
Db 1319 GSCCCCGGCAACAACGCGACCAACAGCAATGAGCAATTCACAGTCTCCCTCCCAAC 1378
QY 1261 CCCAGCTCGATGCAACAACATCTTGATATTTTCAAGAAAGCGATCGAATTTCTGGAAA 1320
Db 1379 CCCAGCTCGATGCAACAACATCTTGATATTTTCAAGAAAGCGAAGTATCTCTGGAGG 1438
QY 1321 GAAAGTGATATATTCCTGTTTATAGATGGGTATTAATACTGAAATGATTTGATCCA 1380
Db 1439 GAAAGTGATATATTCCTGTTTATAGATGGGTATTAATACTGAAATGATTTGATCCA 1498
QY 1381 ATCCCAACACAGAAACCAATCTTTCATGAATCTCCAGCTTTAATACATTCAACTAATGAG 1440
Db 1499 GTTCCAAATACAAAACCAATCTTTCATGAATCTCTAGTTCTGTGCTCGGCTGAGGAA 1558

QY 1441 GAACACACACTCCCGGATATATCAGTTTAACTTTCTCTATTTTCTGTATTAATAATGGA 1500
Db 1559 GATCAACATGCTCCCGCAATATAGTTTAACTTTATCTTTATTTCTTAATAATAATGAG 1618
QY 1501 GATACTGCCTATCTCTGGGAAAAAGAAATGATTTGTGATCAGAGTTGAGGATTTGGAGT 1560
Db 1619 AACACTGCCTATCTCTGGGAAAAAGAAATGAGATGATTTGTGATCAGAGTTAAGATTTGGAGC 1678
QY 1561 GTGAGAGAGCAATTTGGCGGAGGCTTGTGATACCAATTTTGTGATACCAATTTTGTGATGATC 1620
Db 1679 GTTCAGAGAGATGACCTGCGCGAGGCTCAGTTGGATACCGTTTGTGATGATCAGGATTT 1738
QY 1621 GAAGGACTTATCTCTCGGTTTAAATCAAAAAATCAAAAAATCAAAAAATTTGTCTGAGTTGAGG 1680
Db 1739 GAAGGACTTATCTCTCGGTTTAAATCAAAAAATCAAAAAATTTGTCTGAGTTGAGG 1798
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGGTCAACAACCGAGAA 1740
Db 1799 CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGGTCAACAACCGAGAA 1858
QY 1741 AGGACATTTTCTTAAATAGGCAATGCAATTTGAATTTTGTGATACGAGGTTGGGCGGA 1800
Db 1859 AGAATCTTCTTAAATAGGCAATGCAATTTGAATTTTGTGATACGAGGTTGGGCGGA 1918
QY 1801 ACATGCAAGTGTAGGACCTGATTTGTCATAGGATAGAGATCTATCTTAAATAATC 1860
Db 1919 ACATGCAAGTGTAGGACCTGATTTGTCATAGGATAGAGATCTATCTTAAATAATC 1978
QY 1861 TCAGAACAAATCGCAAAATCAGAAAGGATGAAACAAAGGAGGAACTGCTGGGGTCTA 1920
Db 1979 TCAGAGCAATTTGACCAATTTAAAGAGGACGACAAAGAGGGGACTGTTGGGGTCTG 2038
QY 1921 GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCACTCTGCTACTA 1980
Db 2039 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCACTTTTGTCTACTA 2098
QY 1981 TTATCTATAGCTGTCTGATTTGCTCTGCTGATCTGCTGATCTGCTGATCTTCACTAATACT 2040
Db 2099 TTATCTATAGCTGTCTGATTTGCTCTGCTGATCTGCTGATCTTCACTAATACT 2158
QY 2041 GGATGA 2046
Db 2159 GGATGA 2164

RESULT 4
ACC71550
ID ACC71550 standard; DNA; 7778 BP.
XX
AC ACC71550;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6701 (pVR1012-Marburg) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX Laesa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
XX viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX
FN WO2003028632-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030251.
XX
PR 01-OCT-2001; 2001US-0326476P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX

DR	WPI; 2003-371961/35.	Db	2634	AGTAA CGNACGCAACGAATGACACTGGATGTTTCGGGGCTCTTCAAGNATACAAATTCT	2693
XX	New bimodal priming and boosting compositions, useful as viral vaccines, specifically for eliciting an immune response against a filovirus or a disease caused by infection with filovirus.	Qy	661	ACAAGAACCAAAACATGTGTCTCCCTCAAAATACCTCACACTGCCACAGCCGTCGG	720
PT		Db	2694	ACAAGAACCAAAACATGTGTCTCCCTCAAAATACCTCACACTGCCACAGCCGTCGG	2753
XX	Claim 1; Page 185-187; 219pp; English.	Qy	721	GAGATCAAACTCACAAGCACCCCAACTGATGCGACCAACAACTCAATACCAAGGACCCCAAGC	780
XX	The present invention relates to a bimodal priming composition and boosting composition for priming and boosting an immune response to an antigen in an individual. The compositions comprise (a) a priming composition comprised of a DNA plasmid comprising a nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a DNA plasmid selected from Acc71521-Acc71563 and (b) a boosting composition comprised of a replication-deficient adenovirus, comprising a nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus, or influenza virus glycoprotein or nucleoprotein or epitope-bearing domain, or a replication deficient adenovirus selected from constructs Acc71521-Acc71563. The compositions are useful as viral vaccines, specifically for eliciting an immune response against a filovirus or a disease caused by infection with filovirus e.g. Ebola virus infection. The present sequence has a Marburg glycoprotein (gp) open reading frame, Musoke strain, cloned into a VRC6700 backbone	Db	2754	GAGATCAAACTCACAAGCACCCCAACTGATGCGACCAACAACTCAATACCAAGGACCCCAAGC	2813
CC		Qy	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	840
CC		Db	2814	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	2873
CC		Qy	841	ACTTCTGATGGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCCTCCCTCACA	900
CC		Db	2874	ACTTCTGATGGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCCTCCCTCACA	2933
CC		Qy	901	CAACCAAGCAGCCACACAGCAGGAGGAAACAACAACCACTTCCCAAGATGCTGTGACT	960
CC		Db	2934	CAACCAAGCAGCCACACAGCAGGAGGAAACAACAACCACTTCCCAAGATGCTGTGACT	2993
CC		Qy	961	GAACTAGACAAAATAACACAACTGCACAACTGCCATGCCCTCTATAACACTACACACA	1020
CC		Db	2994	GAACTAGACAAAATAACACAACTGCACAACTGCCATGCCCTCTATAACACTACACACA	3053
XX	Sequence 7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;	Qy	1021	ATCTCTACTAAACACACTCTCCAAACACAACTTCAGCAGCTCTCTGCAACCAATTAACAAAC	1080
	Query Match 90.0%; Score 1841.2; DB 10; Length 7778;	Db	3054	ATCTCTACTAAACACACTCTCCAAACACAACTTCAGCAGCTCTCTGCAACCAATTAACAAAC	3113
	Best Local Similarity 93.7%; Pred. No. 0;	Qy	1081	ACCACTATGACACACACAGAGCACAATCACTGAAATAGCAACACAGTCCCTCCCTCG	1140
	Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;	Db	3114	ACCACTATGACACACACAGAGCACAATCACTGAAATAGCAACACAGTCCCTCCCTCG	3173
Qy	1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAAATTCAGGACACAAAATCTCCCC	Qy	1141	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGACGACAAAAA	1200
Db	2034 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAAATTCAGGACACAAAATCTCCCC	Db	3174	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGACGACAAAAA	3233
Qy	61 ATTTTAGATAGCTAGTAATCAACCCCAAAATGGATTCGGTATGCTCCGGAAT	Qy	1201	GGCCCCGCCACCAACGGCACCAACACGACAAATGAGCAGTTTTCACAGTCTCTCCCCACC	1260
Db	2094 ATTTTAGATAGCTAGTAATCAACCCCAAAATGGATTCGGTATGCTCCGGAAT	Db	3234	GGCCCCGCCACCAACGGCACCAACACGACAAATGAGCAGTTTTCACAGTCTCTCCCCACC	3293
Qy	121 CTCAGACAGACAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTGTCT	Qy	1261	CCAGCTCGACTGCAACAACATCTTTGATATATTTTCAGAGAAAGCGATCGATTTCTGGAA	1320
Db	2154 CTCAGACAGACAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTGTCT	Db	3294	CCAGCTCGACTGCAACAACATCTTTGATATATTTTCAGAGAAAGCGATCGATTTCTGGAG	3353
Qy	181 GATTCCTCTTTCGAGGACATCCAAAGCGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT	Qy	1321	GAACTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTTGATCCA	1380
Db	2214 GATTCCTCTTTCGAGGACATCCAAAGCGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT	Db	3354	GAACTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTTGATCCA	3413
Qy	241 GTTGAGTACACAGAGGGGGAGGACCAACATGCTACATATTAAGTGAACGGATCCC	Qy	1381	ATCCCAACACAGAAACAACTTTTGTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
Db	2274 GTTGAGTACACAGAGGGGGAGGACCAACATGCTACATATTAAGTGAACGGATCCC	Db	3414	GTTCCAAAATACAAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCTCGGCTGAGGAA	3473
Qy	301 TCTGGAATAATCCTTGCTGTAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	Qy	1441	GAAACAACACTCCCGCAATATCAGTTTAACTTTCTTTATTTTCTGATTAABAATGGA	1500
Db	2334 TCTGGAATAATCCTTGCTGTAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	Db	3474	GATCAACATGCTCTCCCCCAATTAATAGTTTAACTTTATCTTATTTCTTAAATAATGAG	3533
Qy	361 ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	Qy	1501	GATCTGCTACTCTGGGGAAAAACGAGAAATGATTTGTGATGAGAGTTTCAGAGATTTGGAGT	1560
Db	2394 ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	Db	3534	AACACTGCTACTCTGGGAAAAATGAGAAATGATTTGTGATGAGAGTTTAAAGATTTGGAGC	3593
Qy	421 GGAGCATTTTTTCTGTATGATCGCATTTGCTCCACACAAATGATACCGAGGCAAGTCTTC	Qy	1561	GTGAGGAGGAGCATTTTGGCGGCGAGGCTTAGCTGATACCAATTTTTTTGGCCCTCGAATC	1620
Db	2454 GGAGCATTTTTTCTGTATGATCGCATTTGCTCCACAAATGATACCGAGGCAAGTCTTC	Db	3594	GTTCAGGAGGATGACCTTGGCGCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAT	3653
Qy	481 ACTGAAGGGAACATAGCAGCTATGTTGTCATAAGACAGTGCACAAAATGATTTTCTCG	Qy	1621	GAAGACTCTATCTGCGCGTTTAAATCAAAAATCAGAACAAATTTAGTTTGTAGGTTGAGG	1680
Db	2514 ACTGAAGGGAACATAGCAGCTATGTTGTCATAAGACAGTGCACAAAATGATTTTCTCG	Db	3654	GAAGGACTTTACCTGCTGTTTAAATTAATAATCAAAACAATTTGGTCTGCGAGTTGAGG	3713
Qy	541 CGGCAAGGACAAAGGTTACCGTCATATGAATCTGACTTCTACTTAATAATTTGGCAAGT	Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCTCTTGGAGCTCTTGTAAAGGCTCAACACCGAGGAA	1740
Db	2574 CGGCAAGGACAAAGGTTACCGTCATATGAATCTGACTTCTACTTAATAATTTGGCAAGT	Db	3714	CGTCTAGCCAAATCAAACTGCCAAATCTTTGGAACTCTTATTTGAGAGTCACTACTGAGGNA	3773
Qy	601 ACTAAGGACCAACGATGACACTGATGTTTGGGGCTCTTCAAGNATACAAATTCT				


```

QY 1741 AGGCAATTTCTTAATCAATAGGATGCAATGACTTTTGGCTTACGAGTGGGGCGGA 1800
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3774 AGAATATTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAAGATGGGGAGGA 3833
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 ACATGCAAGTGTAGGACCTGATTGTTCATAGGAATAGAAATCTATCTAAAAATATC 1860
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3834 ACATGCAAGTGTGGACCTGATTGTTCATCGGATAGAAATCTGTCCAAAATATT 3893
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 TCAGACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3894 TCAGACAAATTCACCAATTTAAAAAGGACCAACAAAGAGGAGGACTGGTGGGGTCTG 3953
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1921 GGTGGCAATGGTGGACATCTGACTGGGTGCTTCTCACCATTGGGGCATCTCTACTA 1980
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3954 GGTGGTAAATGGTGACATCCGACTGGGGTGTCTTACTAATCTGGGCATTTTCTCTACTA 4013
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1981 TTATCTATAGCTGTCTGATTGCTCTGTCCTGTATCTGTCTGTATCTTCACTAAATACATT 2040
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4014 TTATCCATAGCTGTCTGATTGCTCTATCCCTGTATTTGTCTGTATCTTACTAAATATATC 4073
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2041 GGATCA 2046
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
ID AA287211 standard; cDNA; 11460 BP.
AC AA287211;
XX
XX 15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX
XX VEE replicon comprising Marburg virus genes.
XX
XX VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV;
KW filovirus; Glycoprotein; GP; nucleoprotein; NP; structural protein; VP40;
KW VP35; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX
XX Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FT CDS 104..2182
FT CDS /product= "Marburg virus nucleoprotein (NP, AAY77128)"
FT CDS /tag= a
FT CDS 2944..3933
FT CDS /product= "Marburg virus structural protein VP35
FT CDS (AAY77130)"
FT CDS /tag= b
FT CDS 4567..5478
FT CDS /product= "Marburg virus structural protein VP40
FT CDS (AAY77129)"
FT CDS /tag= c
FT CDS 5940..7985
FT CDS /product= "Marburg virus glycoprotein (GP, AAY77127)"
FT CDS /tag= d
FT CDS 8864..9697
FT CDS /product= "Marburg virus structural protein VP30
FT CDS (AAY77131)"
FT CDS /tag= e
FT CDS 10200..10961
FT CDS /product= "Marburg virus structural protein VP24
FT CDS (AAY77132)"
FT CDS /tag= f
XX
XX WO200000616-A2.
XX
XX 06-JAN-2000.
PD

```

```

XX 21-JUN-1999; 99WO-US014174.
XX
XX 29-JUN-1998; 98US-0091403P.
XX
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
XX Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;
XX
XX WPI; 2000-160676/14.
XX
XX P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132,
XX AAY77133.
XX
XX Novel Marburg virus vaccines used to induce an immune response against
XX the infection in nonhuman primates.
XX
XX Claim 1; Page 56-57; 57pp; English.
XX
XX The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine
XX of the invention comprises a Venezuelan equine encephalitis virus (VEE)
XX replicon containing copies of the genes encoding Marburg virus (Musoke
XX strain) glycoprotein (GP), a deletion mutant GP (GP-delta-TM),
XX nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24.
XX The Marburg virus genes replace the VEE virus structural protein genes;
XX the result is a self-replicating RNA molecule that encodes its own
XX replicase and transcriptase functions, and in addition makes abundant
XX quantities of the Marburg virus proteins. When replicon RNA is
XX transfected into eukaryotic cells, along with two helper RNAs that
XX express the VEE virus structural proteins, the replicon RNA is packaged
XX into VEE virus-like particles by the VEE virus structural proteins, which
XX are provided in trans. Since the helper RNAs lack packaging signals
XX necessary for further propagation, the resulting VEE replicon particles
XX (VRPs) which are produced are infectious for one cycle but are defective
XX whereby the cell produces the Marburg virus proteins, is ultimately
XX killed by the infection, but does not produce any viral progeny. The VEE
XX replicon provides a potent tool for vaccination with Marburg virus
XX antigens. The replicons, vectors and constructs are used to produce
XX vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic
XX fever) in mammals, to elicit immune responses against Marburg antigens,
XX to confer protective immunity, and to reduce disease symptoms and reduce
XX the severity of disease. Studies of non-human primates vaccinated with
XX the replicon indicate that the vaccine will be efficient in protecting
XX humans against Marburg virus. Prior art Marburg virus vaccination
XX strategies have used formalin-inactivated Marburg virus, which are only
XX partially successful at protecting against Marburg virus infection. The
XX present sequence represents a cDNA corresponding to a VEE virus replicon
XX encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24.
XX (Updated on 15-SEP-2003 to standardise OS field)
XX
XX SQ Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;
Query Match 90.0%; Score 1841.2; DB 3; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5940 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 5999
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 ATTTAGATAGCTAGTAAATCAACCCCAAAATGCGATTCGTTCCGGAAT 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 6000 ATTTAGATAGCTAGTAAATCAACCCCAAAATGCGATTCGTTCCGGAAT 6059
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 CTCAGAAGACAGAAGACGTCCTATCTGATGGATTTCACACTGAGTGGGCAAAAATGTTGCT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 6060 CTCAGAAGACAGAAGACGTCCTATCTGATGGATTTCACACTGAGTGGGCAAAAATGTTGCT 6119
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 GATTTCCTTTGGAGGCATCAAGCGATGGCTTTTTCAGGACAGTGTTACCTCCCAAGAA 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 6120 GATTTCCTTTGGAGGCATCAAGCGATGGCTTTTTCAGGACAGTGTTACCTCCCAAGAA 6179
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACATATATAGTGTAAACGGATCCC 300

```

6180	Db	 GTTGAGTACACAGAGGGGAGGAAGCAAAACATGCTACAAATAAGTGTAAACGGATCCC	6239
301	Qy	 TCTGGAAAAATCCTTGCTGTTTAGATCCTCTACCAACATCCGTGACATATCGGAATGCAAA	360
6240	Db	 TCTGGAAAAATCCTTGCTGTTTAGATCCTCTTACCACAAATCCGTGACTATCTCTAAATGCAAA	6299
361	Qy	 ACTATCCATCATATTCAAAGTCAAAACCCCTCATGCAAGGGATCGCCCTTCATTATATGG	420
6300	Db	 ACTATCCATCATATTCAAAGTCAAAACCCCTCATGCAAGGGATCGCCCTTCATTATATGG	6359
421	Qy	 GGAGCATTTTTTCTGTATGATCGGATTTGCCCTCCACAACAATGTATCGAGCAAAAGTCTTC	480
6360	Db	 GGAGCATTTTTTCTGTATGATCGGATTTGCCCTCCACAACAATGTATCGAGCAAAAGTCTTC	6419
481	Qy	 ACTGAAGGGAACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTTCTCG	540
6420	Db	 ACTGAAGGGAACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTTCTCG	6479
541	Qy	 CGGCAAGGACAAGGGTACCGTTCATATGAATCTGACTTCTACTAATAAATATTCGACAAAGT	600
6480	Db	 CGGCAAGGACAAGGGTACCGTTCATATGAATCTGACTTCTACTAATAAATATTCGACAAAGT	6539
601	Qy	 AGTAAACGGAAACGGAAACGAATGACATGGAATGTTTTCGGCGCTCTTCAAGAATACAAATCT	660
6540	Db	 AGTAAACGGAAACGGAAACGAATGACATGGAATGTTTTCGGCGCTCTTCAAGAATACAAATCT	6599
661	Qy	 ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACACTGCCACAGCCCGCTCG	720
6600	Db	 ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACACTGCCACAGCCCGCTCG	6659
721	Qy	 GAGATCAAACTCAACAGCACCCCAACTGATGCGCACCAAACTCAATACCGAGACCCAAAGC	780
6660	Db	 GAGATCAAACTCAACAGCACCCCAACTGATGCGCACCAAACTCAATACCGAGACCCAAAGC	6719
781	Qy	 AGTGATGATGAGGACTCTCGAAACATCGGCTCAGGGTCCGGAGAACGAGAACCCCAACACA	840
6720	Db	 AGTGATGATGAGGACTCTCGAAACATCGGCTCAGGGTCCGGAGAACGAGAACCCCAACACA	6779
841	Qy	 ACTTCTGATCGGTCACCAAGCAAGGCTTTCATCAACAATGCGCCACTCCCTCACCA	900
6780	Db	 ACTTCTGATCGGTCACCAAGCAAGGCTTTCATCAACAATGCGCCACTCCCTCACCA	6839
901	Qy	 CAACCAAGCACGCCACAGCAGGAGGAAAACAACAACCAATTCCTCCCAAGATGCTGTGACT	960
6840	Db	 CAACCAAGCACGCCACAGCAGGAGGAAAACAACAACCAATTCCTCCCAAGATGCTGTGACT	6899
961	Qy	 GAACCTAGACAAAAATAACAACAACCTGCAACACCGTCCATGCCCCCTCATACACTACCA	1020
6900	Db	 GAACCTAGACAAAAATAACAACAACCTGCAACACCGTCCATGCCCCCTCATACACTACCA	6959
1021	Qy	 ATCTCTACTTAACACACCTCCAAACAACTTCAGCACTCTCTCTGCAACCAATTCACAAAC	1080
6960	Db	 ATCTCTACTTAACACACCTCCAAACAACTTCAGCACTCTCTCTGCAACCAATTCACAAAC	7019
1081	Qy	 ACCAACCAATGACAAACACAGAGCACAAATCACTGAAAATGAGCAAAACCACTGCCCTCG	1140
7020	Db	 ACCAACCAATGACAAACACAGAGCACAAATCACTGAAAATGAGCAAAACCACTGCCCTCG	7079
1141	Qy	 ATAACAAACCTTGCTCCAAACGGGAAATCCCAACCAAGCAAGAGCACCAAGCAAAAAA	1200
7080	Db	 ATAACAAACCTTGCTCCAAACGGGAAATCCCAACCAAGCAAGAGCACCAAGCAAAAAA	7139
1201	Qy	 GGCCCCGCCAACGGGACCAACACAGACAAATGAGCATTTTCAACAGTCTCTCCCCCACC	1260
7140	Db	 GGCCCCGCCAACGGGACCAACACAGACAAATGAGCATTTTCAACAGTCTCTCCCCCACC	7199
1261	Qy	 CCAGCTCGAGTCGACAAACATCTTGATATTTTCAGAGAAAGCGATCGATTTTCTGGAAA	1320
7200	Db	 CCAGCTCGAGTCGACAAACATCTTGATATTTTCAGAGAAAGCGAGTATCTCTCTGGAGG	7259
1321	Qy	 GAAGGTCATATATTCCTCGTTTTTAGATGGGTAAATAAATACTGAAATTTGATTTGATCCA	1380

Db	7260	GAAGGGGACAGTTCCTCTTTCTGGATGGGTTAAATAATGCTCCAATGTGATTTTGACCCCA	73119
Qy	1381	ATCCCAAACACAGAAACAAATCTTTTGATGAATCTCCACAGCTTTAATACTTCAACTAATGAG	1440
Db	7320	GTTCCAATACAAAACAACTTTTGATGAATCTCTAGTTCGTGCTCGCTGAGGAA	7379
Qy	1441	GAACAACACACTCCCGGAATACAGTTTAACTTTCTCTTATTTTCTGTATATAAAATGGA	1500
Db	7380	GATCAACATGCCTCCCAATATAGTTTAACTTTATCTTATTTCTTAATAATAATGAG	7439
Qy	1501	GATACTGCCTACTCTCGGGGAAAACGAGATGATTTGTATGTCAGAGTTGAGGATTTGGAGT	1560
Db	7440	AACACTGCCTACTCTCGGAGAAAATGAGAATGATTTGTATGTCAGAGTTAAGAAATTTGGAGC	7499
Qy	1561	GTGCAGGAGACGATTTGGCGGCAGAGCTTAGCTGGATACCAATTTTTTGGCCCTCGAATC	1620
Db	7500	GTTCAGGAGATGACCTGGCGGCAGGGCTCAGTTGATACCGTTTTTTTGGCCCTGGAAT	7559
Qy	1621	GAAGGACTCTATCTACTGCGGTTTAAATCAAAAAATCAGAACAAATTTAGTTTGATGGTTGAGG	1680
Db	7560	GAAGGACTTTACACTGCTGTTTTTAAATTAATAATCAAAAACAAATTTGGTCTGACAGTTGAGG	7619
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTTAAGGGTCACAAACCGAGGAA	1740
Db	7620	CGTCTAGCCAATCAAACTGCAAAATCCTTGGAACTCTTATTGAGAGTTCACAACTGAGGAA	7679
Qy	1741	AGGACATTTTCCTTAATCAATAGGCATGCAATTTGACTTTTTTGCTTACGAGGTGGGCGGA	1800
Db	7680	AGAACATTTCTCTTAATCAATAGACATGCTATTTGACTTTCTACTCACAGATGGGGAGGA	7739
Qy	1801	ACATGCAAGTGTCTAGSACCTGATTTGTGCATAGGAATAGAAGATCTATCTAAAAATATC	1860
Db	7740	ACATGCAAGTGTCTGGACCTGATTTGTGCATCGGGATAGAAGACTTGTCCAAAAATATT	7799
Qy	1861	TCAGAACAAATCGACAAAATCAGAAAGGATGACAAAAGAGGAGAACTGCTGGGCTCTA	1920
Db	7800	TCAGACAAATTAGCCAAATTAATAAAGAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG	7859
Qy	1921	GGTGGCAATGGTGGACATCTGACTGGGGTGTCTCACCAATTTGGGCATCCTGCTACTA	1980
Db	7860	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGCTACTA	7919
Qy	1981	TTATCTATAGCTGTCTGATTGCTCTCTCTCTGATCTGTCGTATCTTCACTAAATACATT	2040
Db	7920	TTATCCATAGCTGTCTTGATTGCTCTATCTGATTTGTCGTATCTTTACTAAATATATC	7979
Qy	2041	GGATGA 2046	
Db	7980	GGATAA 7985	
RESULT 6			
AAZ51039			
ID	AAZ51039 standard; DNA; 2100 BP.		
XX	AC		
XX	AAZ51039;		
XX	05-JUN-2000 (first entry)		
DT	Marburg virus envelope glycoprotein DNA.		
DE			
DE			
XX			
KW	Marburg virus envelope glycoprotein; viral glycoprotein;		
KW	pseudotyped retrovirus; WMLV; Moloney murine leukaemia virus; antibody;		
KW	transduction; screening agent; immunological agent;		
XX	pharmacological agent; ss.		
XX			
XX	Marburg virus.		
XX			
Key	Location/Qualifiers		
FH	10..2052		
FT	/*cag= a		
FT	/product= "Marburg virus envelope glycoprotein"		
XX			

PN WO200008131-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017702.
XX
XX 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
PR
XX (PURD) PURDUE RES FOUND.
FA
XX
XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
PI Fischbach MA;
PI
XX WPI; 2000-224030/19.
DR
DR P-PSDB; AAY70075.
XX
XX Cells that produce inventive pseudotyped retroviruses having a broad host
PT range useful for introducing nucleotide sequences into target cells.
XX
XX Example 11; Page 60-61; 65pp; English.
PS
XX The present sequence encodes the Marburg virus envelope glycoprotein.
CC Plasmid pMBGPI was produced by cloning into the plasmid pSP72 nucleotide
CC sequences corresponding to nucleotides 5931-8033 from the Marburg virus
CC genome. This is used for generation of stable cell lines transiently
CC producing Marburg-MMLV (Moloney murine leukemia virus) pseudotyped
CC retrovirus with a broad host range. These cells having different viral
CC glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
CC transduced with a desired ribonucleotide sequence. This is used to
CC identify screening agents effective in blocking viral entry into a cell.
CC These agents may be immunological agents like monoclonal or polyclonal
CC antibodies. The pharmacological agents include proteins, peptides or
CC various chemical agents. The pseudotyped retrovirus may be useful in
CC methods of identifying cell surface receptors that allow viral entry
XX
SQ Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;

Query Match 89.2%; Score 1825.2; DB 3; Length 2100;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1915; Conservative 0; Mismatches 128; Indels 3; Gaps 1;

QY 1 ATGAAGACCATGTTTCTTATCAGTCTTATCTTATCTTAAATCAAGGGACAAAAATCTCCC 60
DB 10 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAAATCAAGGGACAAAAATCTCCC 69
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 120
DB 70 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 129
QY 121 CTCCAGAGACAGAGACGTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
DB 130 CTCCAGAGACAGAGACGTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 189
QY 181 GATTCCCTTTGGAGGATCAAGCGATCGAGGATTCAGGACAGGATGTAACCTCCCAAGAT 240
DB 190 GATTCCCTTTGGAGGATCAAGCGATCGAGGATTCAGGACAGGATGTAACCTCCCAAGAT 249
QY 241 GTTGAATACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGAACGGATCCC 300
DB 250 GTTGAATACAGAGGGGAGGAGCCAAACATGCTCAATATAAGTGAACGGATCCC 309
QY 301 TCTGAAAATCTTCTGTTAGATCCTCTCAACACATCCGTAATCCGAAATGCAAA 360
DB 310 TCTGAAAATCTTCTGTTAGATCCTCTCAACACATCCGTAATCCGTAATGCAAA 369
QY 361 ACTATCCATCATATTCAGGATCAAAACCTCATGACAGGGGATCGCTTCATTTATGG 420
DB 370 ACTATCCATCATATTCAGGATCAAAACCTCATGACAGGGGATCGCTTCATTTATGG 429
QY 421 GGAGCATTTTCTGATGATGCTGCTCCACAAATGATGACGAGGCAAGTCTTC 480
DB 430 GGAGCATTTTCTGATGATGCTGCTCCACAAATGATGATGACGAGGCAAGTCTTC 489

QY 481 ACTCAAGGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
DB 490 ---CAAGGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 546
QY 541 CGGCAAGGACAAAGGGTACCGTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 600
DB 547 CGGCAAGGACAAAGGGTACCGTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 606
QY 601 AGTAAACGGAACGCAAAACGAATGACACATCGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 660
DB 607 AGTAAACGGAACGCAAAACGAATGACACATCGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 666
QY 661 ACAAGAAACCAAAACATGTGCTCCGTCGAAATACCTCCACCACTGCGCCACAGCGCGTCCG 720
DB 667 ACAAGAAACCAAAACATGTGCTCCGTCGAAATACCTCCACCACTGCGCCACAGCGCGTCCG 726
QY 721 GAGATCAAACTCACAGCACCCCAACTGATGTCACCACTCAATATACCACGAGCCCAAGC 780
DB 727 GAGATCAAACTCACAGCACCCCAACTGATGTCACCACTCAATATACCACGAGCCCAAGC 786
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
DB 787 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 846
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGTTTTCATCAACAATGCCACCACTCCCTCACA 900
DB 847 ACTTCTGATGCGGTACCAAGCAAGGGTTTTCATCAACAATGCCACCACTCCCTCACA 906
QY 901 CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAGATGCTGTGACT 960
DB 907 CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAGATGCTGTGACT 966
QY 961 GAACTAGACAAAAATAACACAACTGCAACCGTCCATGCGCCCTCATAACTACCACTACCA 1020
DB 967 GAACTAGACAAAAATAACACAACTGCAACCGTCCATGCGCCCTCATAACTACCACTACCA 1026
QY 1021 ATCTCTACTAAACAACCTCCAAACAACAATTTAGAGCACTCTCTCTGCAACCAATTAACAAC 1080
DB 1027 ATCTCTACTAAACAACCTCCAAACAACAATTTAGAGCACTCTCTCTGCAACCAATTAACAAC 1086
QY 1081 ACCCAATGACACACAGCAGCACAATCACTGAAAATGAGCAACCAAGTGCCTCCCTCG 1140
DB 1087 ACCCAATGACACACAGCAGCACAATCACTGAAAATGAGCAACCAAGTGCCTCCCTCG 1146
QY 1141 ATAAACAACCTCTCCAAACGGAATCCCAACACAGCAAGAGAGCAGCAGCAGCAAAAA 1200
DB 1147 ATAAACAACCTCTCCAAACGGAATCCCAACACAGCAAGAGAGCAGCAGCAGCAAAAA 1206
QY 1201 GGCCCGCCCAACAGGCAACCAACAGCAAAATGAGCAATTCACAGTCTCCCTCCCAACC 1260
DB 1207 GGCCCGCCCAACAGGCAACCAACAGCAAAATGAGCAATTCACAGTCTCCCTCCCAACC 1266
QY 1261 CCGAGCTGCACTGCAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAA 1320
DB 1267 CCGAGCTGCACTGCAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAG 1326
QY 1321 GAAGTGATATATTTCCCGTTTATAGTGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
DB 1327 GAAGGCAATGTTCCCTTTTCTGATGGGTTAATAATGCTCCCAATTTGATTTGACCCA 1386
QY 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTTAATTACTCAACTAATGAG 1440
DB 1387 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTGGTCCCTCGGCTGAGAA 1446
QY 1441 GAACACACACTCCCGCAATATCAGTTTAACTTTCTTATTTCTGATGATAAATAATGGA 1500
DB 1447 GATCAACATGCTCCCGCAATATTTAGTTTAACTTTTATCTTATTTCTTAAATAAATGAG 1506
QY 1501 GATCTGCTACTCTGGGGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1507 AACACTGCTACTCTGGGAGAAATGAGATGATGATGATGATGATGATGATGATGATGATG 1566

838 ACTTCTGATGGGTCAACCAAGCAGGGCTTTTCATCAACAATGTCACCCACTCCCTCACCA 897
901 CAACCAAGCAGCCACAGCAGAGGAGAAACACACAAACCAATTCCTCCAGATGCTGTGACT 960
898 CAACCAAGCAGCCACAGCAGAGGAGAAACACAAACCAATTCCTCCAGATGCTGTGACT 957
961 GAACTAGACAAATAAACAACAACCTGACACACCGTCCATGCCCCCTCATACACTACCAACA 1020
958 GAACTAGACAAATAAACAACAACCTGACACACCGTCCATGCCCCCTCATACACTACCAACA 1017
1021 ATCTCTACTAAACAACCTCCAAAACAACCTTCAGCACTCTCTCTGCAACCAATTCACAAAC 1080
1018 ATCTCTACTAAACAACCTCCAAAACAACCTTCAGCACTCTCTCTGCAACCAATTCACAAAC 1077
1081 ACCACCAATGACACACACAGAGCAACATCACTGAAATGAGCAAAACAGTGCCTCCG 1140
1078 ACCACCAATGACACACACAGAGCAACATCACTGAAATGAGCAAAACAGTGCCTCCG 1137
1141 ATAACAACCCCTGCTCCCAACGGGAATCCCAACAGCAAAAGAGCAGCAGCAAAACAA 1200
1138 ATAACAACCCCTGCTCCCAACGGGAATCCCAACAGCAAAAGAGCAGCAGCAAAACAA 1197
1201 GGGCCCGGCACAAACGGCACCACCAACAGCAATGAGCAATTCACCAAGTCCCTCCGCCAC 1260
1198 GGGCCCGGCACAAACGGCACCACCAACAGCAATGAGCAATTCACCAAGTCCCTCCGCCAC 1257
1261 CCCAGCTCGATGCAACAATCTGTATATTTAGAGAAAGCGATCGAATTTCTCGGAAA 1320
1258 CCCAGCTCGATGCAACAATCTGTATATTTAGAGAAAGCGATCGAATTTCTCGGAGG 1317
1321 GAAAGTGATATATTCCTGCTTTAGATGGGTATATAATGCTGAAATGATTTGATATCA 1380
1318 GAAAGCGCATGTTCCCTTTCTGATGGGTATATAATGCTGCAATGATTTGATATCA 1377
1381 ATCCCAACACAGAAACCAATCTTTCATGAATCTCCAGCTTTAATPACTTCAACTAATGAG 1440
1378 GTTCCAAATACAAACCAATCTTTCATGAATCTTCTAGTCTGGTGCTCGCTGAGGAA 1437
1441 GAAACAACATCTCCCGCAATATCAAGTTAACTTTCTTATTTCTGATAAAATGGA 1500
1438 GATCAACATGCTCCCGCAATATTAAGTTAACTTTCTTATTTCTTAAATATAATGAG 1497
1501 GATATGCTCTACTCTGGGCAACAGCAATGATTTGATGATGAGATGAGGATTTGGAGT 1560
1498 AACACTGCTCTACTCTGGGCAACAGCAATGATTTGATGATGAGATTTAAGAAATTTGGAGC 1557
1561 GTGAGGAGGAGCATTTGCGCGCAGGCTTAGCTGATACCAATTTTGGCCCTCGGAATC 1620
1558 GTTCAAGGAGATGACTGCGCGCAGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAT 1617
1621 GAAGGACTCTATPACTGCGGTTTAAATCAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680
1618 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAATTTGCTGCGAGTTGAGG 1677
1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTGTTAAGGCTCAACACGAGAA 1740
1678 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTATTGAGAGTCAACAATGAGAA 1737
1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGACTTTTCTGCTTACAGAGTGGGGCGGA 1800
1738 AGAATCTCTTAAATCAATAGATGCTATTGACTTTTCTACTCAAGATGGGGAGGA 1797
1801 ACATGCAAGGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAATAATATC 1860
1798 ACATGCAAGGCTAGGACCTGATTTGTCATAGGAATAGAGATCTTGTCCAAATAAT 1857
1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920
1858 TCAGAGCAATTTGACCAATTTAAAGGACGAAACAAAGAGGGGACTGGTTGGGGCTG 1917
1921 GGTGCAATGCTGACATCTGCTGGGGTGTCTCACCATTTGCGGATCTGCTACTTA 1980
1918 GGTGTAATGCTGACATCTGCTGGGGTGTCTTACTTAATCTGGGCAATTTGCTACTTA 1977

QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTCTGCTCTGATCTGCTGATCTTCACTAATACATT 2040
DB 1978 TTATCCATAGCTGTTGATTTGCTCTATCTCTGATTTGCTGATCTTACTAATAATATC 2037
QY 2041 GGATGA 2046
DB 2038 GGATAA 2043

RESULT 8
ACC71551

ID ACC71551 standard; DNA; 7005 BP.

XX ACC71551;

XX 10-JUL-2003 (first entry)

XX VRC6702 (pVR1012-x/s Marburg GP (dTM)) plasmid.

XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX Synthetic.

XX WO2003028632-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030251.
XX 01-OCT-2001; 2001US-0326476P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX WPI; 2003-371961/35.
XX New bimodal priming and boosting compositions, useful as viral vaccines,
specifically for eliciting an immune response against a filovirus or a
disease caused by infection with filovirus.
XX Claim 1; Page 188-190; 219pp; English.

The present invention relates to a bimodal priming composition and
boosting composition for priming and boosting an immune response to an
antigen in an individual. The compositions comprise (a) a priming
composition comprised of a DNA plasmid comprising a nucleic acid molecule
encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
composition comprised of a replication-deficient adenovirus, comprising a
nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, or
paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
epitope-bearing domain, or a replication deficient adenovirus selected
from constructs ACC71521-ACC71563. The compositions are useful as viral
vaccines, specifically for eliciting an immune response against a
filovirus or a disease caused by infection with filovirus e.g. Ebola
virus infection. This vector expresses the Marburg virus glycoprotein
without its transmembrane and intracellular domains

Sequence 7005 BP; 1931 A; 1758 C; 1588 G; 1728 T; 0 U; 0 Other;

Query Match 86.4%; Score 1767.6; DB 10; Length 7005;

Best Local Similarity 94.2%; Pred. No. 0;

Matches 1836; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 1 ATGAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAATAATCTCCCC 60

DB 2034 ATGAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAATAATCTCCCC 2093

QY	61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT	120
Db	2094	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT	2153
QY	121	CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTTCACTGAGTGGGCAAAAGTTGCT	180
Db	2154	CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTTCACTGAGTGGGCAAAAGTTGCT	2213
QY	181	GATTCCTCCCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT	240
Db	2214	GATTCCTCCCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT	2273
QY	241	GTGTAGTACACAGAGGGGAGAGCAAAACATGCTACAATPATAGTGTAAACGATCCC	300
Db	2274	GTGTAGTACACAGAGGGGAGAGCAAAACATGCTACAATPATAGTGTAAACGATCCC	2333
QY	301	TCGTGAAATCTCTGCTGTAGATCCTCTACCAACATCCGTGACTATCCGNAATGCAAA	360
Db	2334	TCGTGAAATCTCTGCTGTAGATCCTCTACCAACATCCGTGACTATCCGNAATGCAAA	2393
QY	361	ACTATCCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTATATGG	420
Db	2394	ACTATCCCATCATATTCAAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTATATGG	2453
QY	421	GGAGCATTTTCTGTATGATCGCATTTGCTCCACAAATGTACCGAGGCAAAAGTCTTC	480
Db	2454	GGAGCATTTTCTGTATGATCGCATTTGCTCCACAAATGTACCGAGGCAAAAGTCTTC	2513
QY	481	ACTGAAGGGAACATAGCAGCTATGATTTGTCATATAGACAGTGCACAAAATGATTTTCTCG	540
Db	2514	ACTGAAGGGAACATAGCAGCTATGATTTGTCATATAGACAGTGCACAAAATGATTTTCTCG	2573
QY	541	CGGCAAGGACAGAGGTACCGTCAATATGATCTGACTTCTACTATAAATATTGGACAAGT	600
Db	2574	CGGCAAGGACAGAGGTACCGTCAATATGATCTGACTTCTACTATAAATATTGGACAAGT	2633
QY	601	AGTAAACGGAAACGAAACGAAATGACATCGGATTTTCGGCGCTCTTCAAGAAATACAAATCT	660
Db	2634	AGTAAACGGAAACGAAACGAAATGACATCGGATTTTCGGCGCTCTTCAAGAAATACAAATCT	2693
QY	661	ACAAAGAACCAAAACATGTGCTCGTCCAAATACCTCCACACACTGCCACAGCCCGTCCG	720
Db	2694	ACAAAGAACCAAAACATGTGCTCGTCCAAATACCTCCACACACTGCCACAGCCCGTCCG	2753
QY	721	GAGATCAAACTCACAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGACCCCAAGC	780
Db	2754	GAGATCAAACTCACAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGACCCCAAGC	2813
QY	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGAGAAACGAGAAACCCCAACA	840
Db	2814	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGAGAAACGAGAAACCCCAACA	2873
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA	900
Db	2874	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA	2933
QY	901	CAACCAAGCACGCCACAGCAGAGGAAACACAAACATTTCCCAAGATGCTGTGACT	960
Db	2934	CAACCAAGCACGCCACAGCAGAGGAAACACAAACATTTCCCAAGATGCTGTGACT	2993
QY	961	GAACTAGACAAAATAACACAACTGACCAACCGTCCATGCCCCCTCATAAACAATACCA	1020
Db	2994	GAACTAGACAAAATAACACAACTGACCAACCGTCCATGCCCCCTCATAAACAATACCA	3053
QY	1021	ATCTCTACTAAACACACCTCCAAAACAACTTCAGGACTCTCTCTGACCACTTACAAAAC	1080
Db	3054	ATCTCTACTAAACACACCTCCAAAACAACTTCAGGACTCTCTCTGACCACTTACAAAAC	3113
QY	1081	ACCACCAATGACACACACAGAGCAATCTGAAATATGACAAACAGTGCCTCCCTCG	1140
Db	3114	ACCACCAATGACACACACAGAGCAATCTGAAATATGACAAACAGTGCCTCCCTCG	3173
QY	1141	ATAACAACCTCGCTCCCAACGGGAAATCCCAACACAGCAAGAGAGCAGCAGCAAAAAA	1200

Db	3174	ATAACAACCTCGCTCCCAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA	3233
QY	1201	GGCCCGGCACACAGCGCACCAACACAGCAAAATGAGCAATTTCCAGTCTCTCCCCACCC	1260
Db	3234	GGCCCGGCACACAGCGCACCAACACAGCAAAATGAGCAATTTCCAGTCTCTCCCCACCC	3293
QY	1261	CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAGAGAAAGCGATTCGATTTCTGGAAA	1320
Db	3294	CCCAGCTCGACTGCAACAACTCTTTGTATATTTTCAGAGAAAGCGAAGTATCTCTGGAGG	3353
QY	1321	GAAAGTGATATATTTCCCGTTTTTAGATGGGTTTAAATAATCTGAAATGATTTTGATCCA	1380
Db	3354	GAAAGTGATATATTTCCCGTTTTTAGATGGGTTTAAATAATCTGAAATGATTTTGATCCA	3413
QY	1381	ATCCCAACACACAGAAAAAATCTTTGTATGAAATCTCCAGCTTTTAACTACTTCAATAG	1440
Db	3414	GTTTCAAAATACAAAACAACTTTGTATGAAATCTCTAGTCTGGTGCTCGCTGAGGAA	3473
QY	1441	GAAACAACACTCCCGCGAAATATCAGTTTAACTTTCTCTTATTTTCTGTATATAAATGGA	1500
Db	3474	GATCAACATGCTCTCCCAATATTTAGTTTAACTTTATCTTATTTCTTAATAATAG	3533
QY	1501	GATCTGCTACTCTGGGAAAAACGAGNATGATTTGTATGACAGAGTTGAGATTTGGAGT	1560
Db	3534	AACACTGCTACTCTGGGAAAAATGAGAAATGATTTGTATGACAGAGTTAAGAAATTTGGAGC	3593
QY	1561	GTGAGGAGGACGATTTGGCGGAGGCTTAGCTAGTGATACCATTTTGGCCCTGGAATC	1620
Db	3594	GTTCAGGAGGATGACTGGCGCGAGGCTCAGTTGGATACCGTTTTTGGCCCTGGAATT	3653
QY	1621	GAAGGACTCTATCTACTGCGGTTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGTTGAGG	1680
Db	3654	GAAGGACTTTTACACTGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAA	3713
QY	1681	CGTTAGCTAATCAAACTGCTTAAATCTTTGAGCTCTTTTAAAGGCTCACAACCGAGAA	1740
Db	3714	CGTCTAGCCAATCAAACTGCCAAATCTTTGGAATCTTTTATTTGAGAGTCACTGAGGAA	3773
QY	1741	AGGACATTTTCTTAACTAATAGCATGCAATTCGACTTTTGTCTTACGAGTGGGCGCA	1800
Db	3774	AGAACATTTCTCTTAACTAATAGCATGCTATTGACTTTCTACTCAAGATGGGAGAA	3833
QY	1801	ACATGCAAGGTCTAGGACCTGATTTGTGCAATAGGAAATAGAGATCTATCTAAAAATATC	1860
Db	3834	ACATGCAAGGTCTTGGACCTGATTTGTGCAATCGGATAGAGACTTCTCCAAAAATATT	3893
QY	1861	TCAGAAACAAATCGACAAATTCAGAAAGGATGAAACAAAAGGAGAAACTGGCTGGGCTTA	1920
Db	3894	TCAGAGCAAAATTTGACAAATTTAAAAAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG	3953
QY	1921	GGTGGCAATGGTGGACATCTGACTGGGGT	1950
Db	3954	GGTGGTAAATGGTGGACATCCGACTGGGGT	3983

RESULT 9
ACC71552
ID ACC71552 standard; DNA; 8256 BP.
XX
AC ACC71552;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6710 (pAdapt Marburg GP (dTM)) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
PN WO2003028632-A2.

```
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002MO-US030251.
XX PR 01-OCT-2001; 2001US-0326476P.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX DR WPI; 2003-371961/35.
XX PT New binodal priming and boosting compositions, useful as viral vaccines,
XX PR specifically for eliciting an immune response against a filovirus or a
XX PS disease caused by infection with filovirus.
XX PS Claim 1; Page 190-193; 219pp; English.
XX CC The present invention relates to a binodal priming composition and
XX CC boosting composition for priming and boosting an immune response to an
XX CC antigen in an individual. The compositions comprise (a) a priming
XX CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
XX CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
XX CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
XX CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
XX CC composition comprised of a replication-deficient adenovirus, comprising a
XX CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
XX CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
XX CC epitope-bearing domain, or a replication deficient adenovirus selected
XX CC from constructs ACC71521-ACC71563. The compositions are useful as viral
XX CC vaccines, specifically for eliciting an immune response against a
XX CC filovirus or a disease caused by infection with filovirus e.g. Ebola
XX CC virus infection. The present sequence is an adenovirus shuttle vector,
XX CC expressing the Marburg virus glycoprotein without its transmembrane and
XX CC intracellular domains
XX SQ Sequence 8256 BP; 2048 A; 2049 C; 2108 G; 2051 T; 0 U; 0 Other;

Query Match 86.4%; Score 1767.6; DB 10; Length 8256;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTCTCTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 60
DB 1431 ATGAAGACACATGTTCTCTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 1490

QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 120
DB 1491 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 1550

QY 121 CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 1551 CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 1610

QY 181 GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGACAGAGTGTAACCTCCCAAGAT 240
DB 1611 GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGACAGAGTGTAACCTCCCAAGAT 1670

QY 241 GTTGTAGTACAGAGGGGAGAGCCAAACATGCTACATATAGTGAACGGATCCC 300
DB 1671 GTTGTAGTACAGAGGGGAGAGCCAAACATGCTACATATAGTGAACGGATCCC 1730

QY 301 TCTGGAATCTCTGCTGTAGATCCTCTCAACACATCCGTAATCTCCGAAATGCAAA 360
DB 1731 TCTGGAATCTCTGCTGTAGATCCTCTCAACACATCCGTAATCTCCGAAATGCAAA 1790

QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGAGGGATCGCCCTTCATTTATGG 420
DB 1791 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGAGGGATCGCCCTTCATTTATGG 1850

QY 421 GGAGCATTTTCTGTATGATCGGATTCGCTCCACAAATGTACCGAGGCAAGTCTTTC 480
DB 1851 GGAGCATTTTCTGTATGATCGGATTCGCTCCACAAATGTACCGAGGCAAGTCTTTC 1910

QY 481 ACTGAAGGGAAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAAATGATTTCTCG 540
DB 1911 ACTGAAGGGAAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAAATGATTTCTCG 1970

QY 541 CGGCAAGGACAAAGGGTACCGTTCATATGATCTGATCTTCTACTAATAATATTTGGACAAGT 600
DB 1971 CGGCAAGGACAAAGGGTACCGTTCATATGATCTGATCTTCTACTAATAATATTTGGACAAGT 2030

QY 601 AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 660
DB 2031 AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 2090

QY 661 ACAAGAAACCAACATGTGCTCCGTCGCAAAATACCTCCACACTGCCACAGCCGCTCG 720
DB 2091 ACAAGAAACCAACATGTGCTCCGTCGCAAAATACCTCCACACTGCCACAGCCGCTCG 2150

QY 721 GAGATCAAACTCACAGCACCCCAACTGATGSCACCAAACTCAATACCACGAGCCCAAGC 780
DB 2151 GAGATCAAACTCACAGCACCCCAACTGATGSCACCAAACTCAATACCACGAGCCCAAGC 2210

QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTTCAGGGTCCGGAGAACCGAAGACCCACACA 840
DB 2211 AGTGATGATGAGGACCTCGCAACATCCGGCTTCAGGGTCCGGAGAACCGAAGACCCACACA 2270

QY 841 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACA 900
DB 2271 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACA 2330

QY 901 CAACCAAGCAGCCACAGCAGGAGGAAACAAACAACCAATTCACCAAGATGCTGTGACT 960
DB 2331 CAACCAAGCAGCCACAGCAGGAGGAAACAAACAACCAATTCACCAAGATGCTGTGACT 2390

QY 961 GAACTAGACAAAAATAACAACTGCGCAACCGCTCATGCGCCCTCTATAACACTACCACA 1020
DB 2391 GAACTAGACAAAAATAACAACTGCGCAACCGCTCATGCGCCCTCTATAACACTACCACA 2450

QY 1021 ATCTCTACTAACAAACCTCCAAACAACTTTAGCAGCTCTCTCTGACCAATTCACAAAC 1080
DB 2451 ATCTCTACTAACAAACCTCCAAACAACTTTAGCAGCTCTCTCTGACCAATTCACAAAC 2510

QY 1081 ACCACCAATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140
DB 2511 ACCACCAATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 2570

QY 1141 ATAAACAACTCTCTCCAAACGGGAAATCCCAACAGCAAGCAAGAGCAGCAGCAAGAAAA 1200
DB 2571 ATAAACAACTCTCTCCAAACGGGAAATCCCAACAGCAAGCAAGAGCAGCAGCAAGAAAA 2630

QY 1201 GGCCCCCGCCACAAACGGGCAACCAACAGCAAAATGAGCAATTTCAACAGTCTCCCTCCACC 1260
DB 2631 GGCCCCCGCCACAAACGGGCAACCAACAGCAAAATGAGCAATTTCAACAGTCTCCCTCCACC 2690

QY 1261 CCGAGCTCGACTGACACAACTCTTTGTATATTTTCAGAGAAAGCGATCGATTTCTTGGAAA 1320
DB 2691 CCGAGCTCGACTGACACAACTCTTTGTATATTTTCAGAGAAAGCGAGTATCCTCTGAGG 2750

QY 1321 GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
DB 2751 GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 2810

QY 1381 ATCCCAACACAGAAACAACTTTTGATGAATCTCCAGCTTTTAATTAATCTCAACTAATGAG 1440
DB 2811 GTTCCAAATACAAAACAACTCTTTGATGAATCTCTAGTCTCTGGTGCTCGCTCGAGGAA 2870

QY 1441 GAAACAACTCTCCCGCAATATCAGTTTAATCTTTCTTATTTCTTGATTAATAATGGA 1500
DB 2871 GATCAACATGCTCCCGCAATATAGTTTAATCTTTCTTATTTCTTAATAATAATGAG 2930

QY 1501 GATCTGCTACTCTGGGGGAAACAGAGATGATTTGTGATGAGAGTTTCAGAGTTTGGAGT 1560
DB 2931 AACATGCTACTCTGGGAGAAATGAGATGATTTGTGATGAGAGTTTGGAGT 2990
```



```
QY 1561 GTGAGGAGGACGATTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
DB 2991 GTTCAGGAGGATGACTGGCGCAGGGCTCACTTGGATACCGTTTGGCCCTGGAAT 3050
QY 1621 GNAGGACTCTATACGCGCGTTTAAATCAAAAATCAGAACATTTAGTTGAGTTGAGG 1680
DB 3051 GAAGGACTTTACACTGCTGCTTTTAAATTAATAAATCAAAAATTTGGTCTGCAGGTTGAGG 3110
QY 1681 CGCTTAGCTAAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGCTCACAACCCAGGAA 1740
DB 3111 CGTCTAGGCAATCAAACTGCTAAATCCTTGGAGCTCTTATTGAGAGTCAAACTGAGGAA 3170
QY 1741 AGGACATTTCTTAAATCAATAGGATGCAATTTGACTTTTTCCTTACGAGGTGGGGCGGA 1800
DB 3171 AGAACATTTCTTAAATCAATAGATGCTATTGACTTTCTACTCACAAGATGGGGGGA 3230
QY 1801 ACATCGAAGTCTAGGACCTGATTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
DB 3231 ACATGCAAAAGTCTTGGACCTGATTGTTGTCATCGGGATAGAGACTTGTCCAAAAATAT 3290
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
DB 3291 TCAGAGCAAAATGACAAATTAATAAGGACGAAACAAAAGAGGGACTGTTGGGGTCTG 3350
QY 1921 GGTGCAAAATGTTGACATCTGACTGGGGT 1950
DB 3351 GGTGGTAAATGTTGACATCCGACTGGGGT 3380

RESULT 10
ADM48345
ID ADM48345 standard; DNA; 19112 BP.
AC ADM48345;
XX
DT 18-NOV-2004 (first entry)
DE Marburg virus viral protein genomic DNA.
DE
XX
KW Filovirus; viral protein; NP protein; VP30 protein; VP35 protein;
KW VP40 protein; L protein; membrane-associated protein; matrix protein;
KW polymerase complex protein; minor nucleoprotein; glycoprotein; GP;
KW viral mutagenesis study; vaccine; gene therapy; gene; ds.
OS
XX
FH Marburg virus.
FT
FT Key Location/Qualifiers
FT CDS 103..2190
FT /tag= a
FT /product= "NP protein"
FT CDS 2944..3933
FT /tag= b
FT /product= "VP35 protein"
FT CDS 4567..5478
FT /tag= c
FT /product= "VP40 protein"
FT CDS 5940..7985
FT /tag= d
FT /product= "GP protein"
FT CDS 8867..9712
FT /tag= e
FT /product= "VP30 protein"
FT CDS 10205..10966
FT /tag= f
FT /product= "VP24 protein"
FT CDS 11479..18474
FT /tag= g
FT /product= "L protein"
XX
PN US2003215794-A1.
XX
PD 20-NOV-2003.
```

```
XX 29-JAN-2003; 2003US-00353856.
XX 31-JAN-2002; 2002US-0353972P.
XX (KAWA/) KAWAKA Y.
XX PA (JASE/) JASENOSKY L D.
XX PA (NEUM/) NEUMANN G.
XX Kawaoka Y, Jasenovsky LD, Neumann G;
XX WPI; 2004-010776/01.
XX P-PSDB; ADM48338, ADM48339, ADM48340, ADM48341, ADM48342, ADM48343,
XX ADM48344.
XX Preparing filovirus useful in producing vaccines and gene therapy
XX vectors, comprises contacting a cell with a vector comprising a promoter
XX operably linked to a filovirus genomic cDNA or protein.
XX Disclosure; SEQ ID NO 27; 104pp; English.
XX The invention relates to a method of preparing filovirus. The method
XX involves contacting a cell with a vector comprising a promoter operably
XX linked to a filovirus genomic cDNA or its portion linked to a
XX transcription termination sequence or to a DNA segment encoding a
XX filovirus RNA transcriptase-polymerase, filovirus NP, VP30 or VP35, to
XX yield infectious filovirus. The portion of the cDNA when transcribed
XX yields a RNA capable of being packaged into filovirus virions or capable
XX of being replicated in the presence of filovirus proteins. The method is
XX useful for producing filovirus useful in viral mutagenesis studies and in
XX the production of vaccines and gene therapy vectors. The present sequence
XX is marburg virus (NP; L; polymerase complex-VP35; matrix protein-VP40;
XX glycoprotein-GP; minor nucleoprotein-VP30; membrane-associated protein-
XX VP24) genomic DNA.
XX
XX Sequence 19112 BP; 6176 A; 3757 C; 3578 G; 5601 T; 0 U; 0 Other;
XX
Query Match 82.7%; Score 1692.4; DB 13; Length 19112;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTCTTATCAGTCTTATTAATCAAGGACAAAAATCTCC 60
DB 5940 ATGAAGACCACATGCTCTTTATCAGTCTTATTAATCAAGGATAAAACTCTCC 5999
QY 61 ATTTTAGAGTAGCTAGTAATAATCAACCCAAATGCGATTGCGTATGTCGGA 120
DB 6000 ATTTTAGAGTAGCTAGTAATAATCAACCCAAATGCGATTGCGTATGTCGGA 6059
QY 121 CTCAGAGACAGAGACGTCCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCAGAGACAGAGAGATGTCCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCTTTGGAGGCATCCAGGATGGGCTTTCAGGACAGGTGTAACCTCCCAAGAT 240
DB 6120 GATTCCCTTTGGAGGCATCCAGGATGGGCTTTCAGGACAGGTGTAACCTCCCAAGAT 6179
QY 241 GTTGAGTACAGAGGGGGAGGCCAAACATGCTACAATATAAGTGTACGGATCC 300
DB 6180 GTTGAGTATACAGAGGGGGAGGCCAAACATGCTACAATATAAGTGTACGGATCC 6239
QY 301 TCTGAAAATCCTTGCTTTAGATCCTCTACCAACATCCGTAATCCGAAATGCAAA 360
DB 6240 TCTGAAAATCCTTGCTTTAGATCCTCTACCAACATCCGTAATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 420
DB 6300 ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 6359
QY 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAGTCTTC 480
DB 6360 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAGTCTTC 6419
```


QY 481 ACTGAGGGAACATAGCAGCTATGATTCTCAATAAGACAGTGCACAAAATGATTTCCTCG 540
Db 6420 ACTGAAGGGAACATAGCAGCTATGATTCTCAATAAGACAGTGCACAAAATGATTTCCTCG 6479
QY 541 CGGCAAGGACAGGAGTACCGTCAATGAACTCTGACTTCTACTAAATAAATATTTGGACAAGT 600
Db 6480 AGGCAAGGACAGGAGTACCGTCAATGAACTCTGACTTCTACTAAATAAATATTTGGACAAGT 6539
QY 601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTTCGGGCTCTTCAAGAATATCAAAATCT 660
Db 6540 AACAAATGGAACACAAACGAATGACACTGGATGTTTTCGGGCTCTTCAAGAATATCAAAATCT 6599
QY 661 ACAAGACACCAACATGCTCCGTCACCAATACCTCCACCACTGCCACACAGCCGCTCG 720
Db 6600 ACGAAGATCAAAATGCTCCGTCACCAATACCTCCACCACTGCCACACAGCCGCTCG 6659
QY 721 GAGATCAAACTCACAAAGCACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 780
Db 6660 GAGATCAAACTCACAAAGCACCCCACTGATGCCACCACTCAATACCAAGCCCAAGC 6719
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGTCCAGGTCGGAGAACGAGAACCCCAACA 840
Db 6720 AATGATGATGAGGACCTCATAACTCCGGTCCAGGTCGGAGAACGAGAACCCCAACA 6779
QY 841 ACTTCTGATGCGGTCAACGACGAGGCTTTTCATCAACATGCCACCACTCCCTCACC 900
Db 6780 ACTTCTGATGCGGTCAACGACGAGGCTTTTCATCAACATGCCACCACTCCCTCACC 6839
QY 901 CAACCAAGCAGCCACACAGGAGGAAACACACAAACCACTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCACACAGGAGGAAACACACAAACCACTTCCCAAGATGCTGTGACT 6899
QY 961 GAACTAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCTCATACCACTACCA 1020
Db 6900 GAACTAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCTCATACCACTACCA 6959
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGACCACTTACAAAC 1080
Db 6960 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGATATCACTACAAAC 7019
QY 1081 ACCACATGACACACACAGACCAATCACTGAAATGAGCAACCAAGTCCGCTCCCTCG 1140
Db 7020 ACCACATGACACACACAGACCAATCACTGAAATGAGCAACCAAGTCCGCTCCCTCG 7079
QY 1141 ATAAACACCTGCTCCCAACGCGAAATCCCAACACAGCAAGAGACCAAGCAGCAAAAA 1200
Db 7080 AAAACACCTGCTCCCAACGCGAAATCTTACCAACAGCAAGAGCACTTAAACACAGAA 7139
QY 1201 GGGCCCGCACAAAGGACCAACAGCAATGAGCAATTTTCAACAGTCTCCCTCCCAACC 1260
Db 7140 GGGCCCGCACAAAGGACCAATATGACAAATGGGCAATTTTAAACAGTCTCCCTCCCAACC 7199
QY 1261 CCCAGCTGACGACACATCTGTATATTTTCAAGAAAGCGATCGATTTCTCGGAA 1320
Db 7200 CCCAACCAGCACACATCTGTATATTTTCAAGAAAGCAAGTATCTCTCGGAGG 7259
QY 1321 GAAAGTGATATATCCCGTTTTTGTAGTGGTTTAAATAAATACTGAAATGATTTTGATCCA 1380
Db 7260 GAAAGTGATATATCCCGTTTTTGTAGTGGTTTAAATAAATACTGAAATGATTTTGATCCA 7319
QY 1381 ATCCCAACACAGAAACAACTTTGATGAATCTCCAGCTTTAATACCTCACTAATGAG 1440
Db 7320 GTTCCAAATACAAAGACGATCTTTGATGAATCTTCTAGTCTGCTCGGCTGAGAA 7379
QY 1441 GAAACACACCTCCCGCAATATGATTTTCTTCTATTTTCTGATTAATAAATGGA 1500
Db 7380 GATCAACATGCTCCCGCAATATGATTTTCTTCTATTTTCTGATTAATAAATGGA 7439
QY 1501 GATCTGCTACTCTGGGGAACACGAGATGATTTGTGATGACAGTTCAGGATTTGGAGT 1560
Db 7440 AACCTGCTACTCTGGGGAACATGAGACGATTTGTGATGACAGTTCAGGATTTGGAGT 7499
QY 1561 GTGACAGGAGGACGATTTTGGCGGACGGCTTAGCTGATACCAATTTTTTTGGCCCTCGAATC 1620

Db 7500 GTTCAGGAGGATGACCTGGCAGAGGGCTCAGTTTGGATACCGTTTTTTGGCCCTCGAATC 7559
QY 1621 GAAGGACTCTATCTGCTGCGGTTTAAATCAAAAATCAGAACAAATTTAGTTTGTAGTTGAGG 1680
Db 7560 GAAGGACTTTATCTGCTGCTGTTTAAATTAATAAACCAAAACAATTTGGTCTGCAGGTTGAGG 7619
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTTTAAAGGTCACAAACGAGAA 1740
Db 7620 CGTCTAGCCATCAAACTGCTCAAAATCTCTTGGAACTCTTTTAAAGAGTCACAAACGAGAA 7679
QY 1741 AGGACATTTCTCTTAATCAATAGCATGCAATTCGATTTTGTCTTACGAGGTGGGCGGA 1800
Db 7680 AGGACATTTCTCTTAATTAATAGACATGCCATTCGATTTTCTCTCAAGGTGGGAGGA 7739
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTTGTGCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 7740 ACATGCAAGGTGCTTGGACCTGATTTTGTGCATTTGGAATAGAGACTTCTCCAGGAATAT 7799
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGAGGAAACTGGCTGGGCTCTA 1920
Db 7800 TCGGAACAAATTTGACCAATCAAAAAGATGAACAAAGAGGAGGACTGGTTGGGCTCTA 7859
QY 1921 GGTGCAAAATGCTGACATCTGCTGGGCTGTTCTCAACCAATTTGGGCACTCTGCTACTA 1980
Db 7860 GGTGTAATGCTGACATCTGCTGGGCTGTTCTTACTAACTTGGGCAATTTTGTCTACTA 7919
QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGTATCTGCTGATCTTCACTAAATACAT 2040
Db 7920 TTATCCATAGCTGTTGATTTGCTCTATCTGATTTGCTGATTTTACCAATATATC 7979
QY 2041 GGATGA 2046
Db 7980 GGGTAA 7985

RESULT 11
ABT13456
ID ABT13456 standard; DNA; 2046 BP.
XX
AC ABT13456;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein RVN-GP1/GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
DR P-FSDB; ABU18478.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Disclosure; Page 89-91; 94pp; English.
XX

CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
XX

SQ Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;

Query Match 74.2%; Score 1518; DB 10; Length 2046;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 1716; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

```
QY 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCGGAACT 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GTTTTAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTGAGTGTCTCCGGAACC 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CTCGAGACAGACAGAGCTTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CTCCAAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GATTCCTCTTGGAGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GATTCCTCTTGGAGCATCTAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GTTGTAGTACACAGAGGGGAGGAGCAAAACATGCTACAATATAGTGTAAACGGATCCC 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GTTGTAGTATACGAGAGGAGGAGCAAAACATGTTTACAATATAGTGTAAACGACCT 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 TCTGGAATCTTCTGCTGTAGATCCTCTCAACATCCGTGACTATCCGNAATGCAAA 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 TCTGGAATCTTCTGCTGTAGATCCTCTCAACATCCGTGACTATCCGNAATGTAATA 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 ACTGTTTCATCATATCAAGGTCAAAACCTCATGACAGGGGATTCGCTCATTGTTGG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAATGTACCGAGGCAAGTCTTC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GGCGCATTTTTCTGTATGATCGGTGCTCTCAACAATGTACCGAGGCAAGTCTTC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCCAGAAATGATTTTCT 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCCAGAAATGATTTTCT 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 CGGCAAGGACAAAGGTACCGTCAATGATCTGACTTCTACTAATAAATATTGGACAAGT 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 AGGCAAGGACAAAGGTATCGTCACATGAATCTGACTTCCACCAATATATTGGACAAG 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 AGTAAACGGAACCAACGAAATGATCGGATGTTTGGCGCTCTTCAAGAAATCAATTTCT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 AGCAATGAACGCAAGAAATGATCGGATGTTTGGCATCTCCAGAAATCAAACTCC 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 ACAAGAACCAACATGCTCGTCCCAAAATACCTCCACCATGCGCCACAGCCGCTCG 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 ACAAAACAATCAAAACATGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 GAGATCAAACTCACAAGACCCCAACTGATGCCAACCACTCAATACACGAGCCCAAGC 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 721 AGCATTCATCTACAATATCTCAATTAATATCTGCTAATCTGGAATCTATGAACCAAGT 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 AGTGTATGATGAGACCTCGCAATCCGGCTCAGGGTCGAGGAACGGAACCCCAACACA 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 781 AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 ACTTCTGATGCGGTACCAAGCAGGCGCTTTCATCAAGATGCCACCACTCCCTCACC 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTTGTCCACTCTCTACT 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 901 CAACCAAGCAGCGCCACAGCAAGAGGAAACAAACAACCAATCCCAGATGCTGTGACT 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 901 CATCAAGCACCTCAACAATGAGCAAAAACAGTAGCAATCTTCCCGACATGCTGTAACT 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 GAACCTAGACAAAATAACAACTGCAACACGGTCCATGCCCCCTCAACACTACCAACA 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 961 GAGCACAATGGAAACCGACCCCAACAACAACACGAAACGCTCCCAACANTACTAATA 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 ATCTCTACTAAACAACCTCCAAAACAACAACCTTACAGCACTCTCTGCAACCAATTA 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1021 ACTCCACCTTAACACTCTCAAGTACAACTCAGTACTCTTCCCCTCCAAACCGCAAC 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1081 ACCACCAATGACAAACACAGAGCACAAATCACTGAAATAGCAAAACAGTGCCCTCG 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1081 ATCAACCAATATGATACAAACGCTGAACCTAGCAGAAAGCAACAACCAATGCTCAG 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1141 ATAAACACCTGCTCCAAACGGGAATCCCAACACGAAAGACACAGCAAGCAAAA 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGCAAGACACCAACAGCAAC 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1201 GGCCCGCCACAAACGGCACCAACAACAGCAAAATGAGCAATTTCCAGCTCTCTCCC 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1201 AACATCATGACGACATCAGATATACAGCAAAACACCCCAACAATTTCTTCCG 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1261 CCCAGCTCGACTGCACAAACATCTTTGTATATTTCAAGAAAGCGATCGATTTCTG 1320
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1261 TCTAGTCGACAAACCGCCCTCTATATCTTTAGAAAGAAACGATCGATTTCTTGAA 1320
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1321 GAAGGTGATATATTTCCCGTTTTAGATGGGTTAATAACTGAAATTTGATTTGATCCA 1380
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1321 GAAGGTGATATATTTCCCGTTTTAGATGGGTTAATAACTGAAATTTGATTTGATCCA 1380
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAAT 1440
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAAT 1440
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 GAACAACAACCTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA 1500
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 GAACAACAACCTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA 1500
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 GATCTGCTACTCTGGGAAAAACGAGATGATTTGTGATGACAGTTGAGGATTTGGAGT 1560
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 GATCTGCTACTCTGGGAAAAACGAGATGATTTGTGATGACAGTTGAGGATTTGGAGT 1560
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 GTGAGGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1561 GTGAGGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1621 GAAGGACTCTATCTGCGGTTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGTTG 1680
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1621 GAAGGACTCTATCTGCGGTTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGTTG 1680
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTTTAAAGGTCACACCCAGG 1740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTTTAAAGGTCACACCCAGG 1740
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1741 AGGACATTTTCTTAACTAGGATCAATTTGACTTTTGTCTTACGAGGTGGGCGGA 1800
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1741 AGGACATTTTCTTAACTAGGATCAATTTGACTTTTGTCTTACGAGGTGGGCGGA 1800
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGATAGAAATAGAAATCTATCTAAAAAT 1860
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGATAGAAATAGAAATCTATCTAAAAAT 1860
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 TCAGAAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAACTGGCTGGGCTCT 1920
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1861 TCAGAAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAACTGGCTGGGCTCT 1920
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1921 GGTGCAAAATGTTGACATCTGACTGGGTGTTCTACCAATTTGGGATCTCTGTACT 1980
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1921 GGTGCAAAATGTTGACATCTGACTGGGTGTTCTACCAATTTGGGATCTCTGTACT 1980
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
QY 1141 ATAAACACCTCGCTCCACAGGGAATCCACACAGCAAGACGACCAGGCAAAAA 1200
Db 1237 AACACAACTCTAGATCCACACAGAAATCCACACAGGACAGACACCAACGCAAC 1296
QY 1201 GCGCCGCCACAAACGCGCACAAACAGCAAAATAGCATTTCCACAGTCTCCGCCACC 1260
Db 1297 ACATCATCATGACGACATCAGATATACAGCAACACCCACAAATCTCTCCGAT 1356
QY 1261 CCCAGCTCGACTGCACAAACATCTTGATATATTTAGAGAAAGCGATCGATTTCTGAAA 1320
Db 1357 TCTAGTCCGACAAACCGCCCTCCTATATCTTTAGAAAGAAACGAAGCATTTCTCGAAA 1416
QY 1321 GAAGTGATATATTCCTGTTTTAGATGGTTTAATAAATAGTGAATTTGATGCA 1380
Db 1417 GAAGTGATATATTCCTGTTTTAGATGGTTTAATAAATAGTGAATTTGATGCA 1476
QY 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATTAATCTCAACTAATGAG 1440
Db 1477 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATTAATCTCAACTAATGAG 1536
QY 1441 GAACACACACTCCCCCGAATATCATGTTTAACTTTCTTTATTTTCTGATAAAAAATGGA 1500
Db 1537 GAACACACACTCCCCCGAATATCATGTTTAACTTTCTTTATTTTCTGATAAAAAATGGA 1596
QY 1501 GATAGTCTACTCTGGGAAAAACGAGAAATGATGTGATGACAGATTGAGGATTTGGAGT 1560
Db 1597 GATAGTCTACTCTGGGAAAAACGAGAAATGATGTGATGACAGATTGAGGATTTGGAGT 1656
QY 1561 GTGAGGAGGACGATTGTGGCGCAGGCTTACGTGATACCAATTTTGTGGCCCTGGAATC 1620
Db 1657 GTGAGGAGGACGATTGTGGCGCAGGCTTACGTGATACCAATTTTGTGGCCCTGGAATC 1716
QY 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Db 1717 GAAGGACTCTATCTGCGGTTTAAATCAAAAAATCAGAAATTTAGTTTGTAGTTGAGG 1776
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTGTTAAGGTCACAAACGAGAA 1740
Db 1777 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTGTTAAGGTCACAAACGAGAA 1836
QY 1741 AGGACATTTTCTTAATCAATAGGACATGCAATTTGCTTACGAGGTGGGCGGA 1800
Db 1837 AGGACATTTTCTTAATCAATAGGACATGCAATTTGCTTACGAGGTGGGCGGA 1896
QY 1801 ACATGCAAGTGCTAGGACCTGATTGTTGCATAGGAATAGAAATCTATCTAAAAATATC 1860
Db 1897 ACATGCAAGTGCTAGGACCTGATTGTTGCATAGGAATAGAAATCTATCTAAAAATATC 1956
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 1957 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGGTCTA 2016
QY 1921 GGTGGCAATGTGGACATCTGACCTGGGTGTTCTCACCATTTGGGCATCTGCTACTA 1980
Db 2017 GGTGGCAATGTGGACATCTGACCTGGGTGTTCTCACCATTTGGGCATCTGCTACTA 2076
QY 1981 TTATCTATAGCTGTTCTGATGCTCTGCTGCTGATCTGTCGATCTTCACTAAATACATT 2040
Db 2077 TTATCTATAGCTGTTCTGATGCTCTGCTGCTGATCTGTCGATCTTCACTAAATACATT 2136
QY 2041 GGATGA 2046
Db 2137 GGATGA 2142
```

RESULT 13

ABT13453

ID ABT13453 standard; DNA; 2046 BP.

XX AC

XX ABT13453;

XX DT 30-JAN-2003 (first entry)

XX

DE DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

KW immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

OS Chimeric.

XX WO200279239-A2.

PN 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.

XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX Grogan CC, Hevey MC, Schmaljohn AL;

XX WPI; 2003-040651/03.

XX P-PSDB; ABJ18475.

PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,

PT useful for inducing an immune response against infection of different

PT filoviruses, specifically against both Ebola and Marburg viruses.

XX Claim 18; Page 78-80; 94pp; English.

XX The invention relates to a chimeric filovirus glycoprotein (GP) protein

XX comprising GP1 and GP2, where GP1 is from a filovirus different than that

XX of GP2. The chimeric filovirus GP protein is useful for inducing an

XX immune response against infection of different filoviruses, specifically

XX against both Ebola and Marburg viruses by being used as a vaccine. This

XX polynucleotide sequence represents a DNA encoding a chimeric filovirus

XX protein of the invention

XX Sequence 2046 BP; 662 A; 486 C; 379 G; 519 T; 0 U; 0 Other;

SQ Query Match 64.5%; Score 1319.6; DB 10; Length 2046;

Beat Local Similarity 77.8%; Pred. No. 0;

Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTCTCTTATCAGTCTTATCTTAATCAAGGACAAAAATCTCCC 60

Db 1 ATGAAGACCACATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAACTCTCCT 60

QY 61 ATTTTAGAGTAGCTAGTAAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAC 120

Db 61 GTTTTAGAAATTTGCTAGTAAACAGCCAACTCAAGATGTAGATTGAGTGTCTCCGGAAC 120

QY 121 CTCAGAAACAGAAAGACGCTCCATCTGATGGGATTCACACTGAGTGGCAAAAAAGTTGCT 180

Db 121 CTCAGAAACAGAAAGATGTTTCATCTGATGGGATTCACACTGAGTGGCAAAAAAGTTGCT 180

QY 181 GATTCCCTTTGGAGGACATCAACGATGCGCTTTACAGACAGGTGTACTCTCCCAAGAA 240

Db 181 GATTCCCTTTGGAGGACATCAACGATGCGCTTTACAGACAGGTGTACTCTCCCAAGAA 240

QY 241 GTTCAGTACAGAGGGGAGAGCCAAACATGCTACAAATAGTAAAGTAAACGATCCC 300

Db 241 GTTCAGTATACGGAAGGAGAGCCAAACATGTTTACAATATAGTAAACGATCCC 300

QY 301 TCTGAAAAATCTTGTGTTAGATCTCTCTACCAACATCCCGTACTATCCGAAATGCAAA 360

Db 301 TCTGAAAAATCTTGTGTTAGATCTCTCTACCAACATCCCGTACTATCCGAAATGCAAA 360

QY 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGACAGGGGATCGCCCTTCATTATGG 420

Db 361 ACTGTTCATCATATTTCAAGGTCAAAACCCCTCATGACAGGGGATCGCCCTTCATTATGG 420

QY 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGAAAGTCTTC 480

Db 421 GGGGCAATTTTCTTGATGTCGGTTGCTCTACAACAATGTACCGAGCAAGTCTTC 480
Qy 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTGTAAATAGACAGTTCACAGAATGATTTTCT 540
Qy 541 CGGCAAGCAGCAGGTCACCGTCATATGAATCTGACTTCTACTAATAATATTTGACAGT 600
Db 541 AGGCAGCAGCAGGTTATCGTCACATGAACCTTGACTCCCAATAATAATATTTGACAGC 600
Qy 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTTGCGGCTCTTTCAAGAAATACAATTTCT 660
Db 601 AGCAATGAACGCGAGAGAAATGATAGCGGATGTTTTGGCATCTCCCAAGAAATACAATCTC 660
Qy 661 ACAAGAACCAAAACATGTGCTCCGTCGCAAAATACCTCCACACTGCCACAGCCGCTCG 720
Db 661 ACAAAACATCAACATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
Qy 721 GAGATCAAACTCACAGCACCCCAACTGATGCCAACCAACTCAATACCAAGGACCAAGC 780
Db 721 AGCATTTCACTCTACAATACTCAAAATTAATCTGCTAAATCTGGAACATATGAACCCAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGCTCAGGCTCCGAGAACGAGAACCCCAACACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGAACAGGGGCCCAACACA 840
Qy 841 ACTTCTGATGGGTCAACGAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGCTCCACTCTCACTA 900
Qy 901 CAACCAAGCAGCCACAGCAGGAGGAGAACACACAAACCAATCCCAAGATGCTGTACT 960
Db 901 CATCCAAGCACTCACACATGAGCAAAACAGTAGTACGAATCTTCCCGACATGCTGTAACT 960
Qy 961 GAACTAGACAAAATAACAACTGCACAACTCGTCCATGCCCTCCATATAACACTACCACA 1020
Db 961 GAGCAATGAGAACCGACCCACACACAAACAGCAGCAACGCTCTCCACAACTACTAATA 1020
Qy 1021 ATCTCTACTAACAACCTCCAAACACAACTTTCAGACTCTCTCTGACCACTTACAAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
Qy 1081 ACCCAATGACACACAGCAGCACATCACTGAAATGAGCAAAACGAGTCCGCTCG 1140
Db 1081 ATCCAAATATGATACACACGTAACGTAAGAGAAAGCGAACAAACCAATGCTCAGTTG 1140
Qy 1141 ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGCAGCAAAAGAGCACAGCAAAAAA 1200
Db 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGCAGCAAGACACCAAGCACACACC 1200
Qy 1201 GGCCCCGCCAACAGGCCACCAACACGCAATGAGCATTTTCAACAGTCTCTCCCCCACC 1260
Db 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAACAAATTTCTCTCGAT 1260
Qy 1261 CCCAGTCTGACACACATCTGTATATTTTCAAGAAAGCGATCGATTTCTTGAA 1320
Db 1261 TCTAGTCCGCAACCCGCTCTCTATATCTTGAAGAAAGAACCGATCGATCTCTGGAG 1320
Qy 1321 GAAGTGATATATTTCCGTTTTTACATGGGTTTATAATAATCTGAAATGATTTTGTATCCA 1380
Db 1321 GAAGCGACATGTTCCCTTTCTGATGGGTTTATAAATGCTCCAAATGATTTTGTACCCA 1380
Qy 1381 ATCCAAACACAGAAACAACTTTTGTATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAAAACAACTTTTGTATGAATCTTCTAGTTCTGTGCTCGCTCGAGAA 1440
Qy 1441 GAAACACACATCCCGGAAATATCAGTTTAACTTTCTTATTTTCTGTGATAAAATGGA 1500
Db 1441 GATCAACATGCTCCCGGAAATATAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Qy 1501 GATCTGCTACTCTGGGGAACCAAGAAATGATTTGTGATGACAGATTTGAGGAT 1560
Db 1501 AACCTGCTACTCTGGGGAACCAAGAAATGATTTGTGATGACAGATTTAAGATTTGGAGC 1560

Qy 1561 GTGCAGAGACAGATTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
Db 1561 GTTCAGAGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTTTGGCCCTGSAAT 1620
Qy 1621 GAAGGACTCTATCTGCGGGTTTAAATCAAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG 1680
Db 1621 GAAGGACTTTACTGCTGCTGTTTTAATTAANAATCAAAAACAATTTGGTCTGCAGGTTGAG 1680
Qy 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTTTGTTAAGGGTCACAAACGAGGAA 1740
Db 1681 CGCTAGCCCAATCAAACTGCAAAATCTTGGAACTCTTATTTAGAGATCACAACTGAGGAA 1740
Qy 1741 AGGACATTTCTTAAATCAATAGCATGCTATTTGACTTTTCTACTCACAAGATGGGGGGA 1800
Db 1741 AGAATCTTCTTAAATCAATAGCATGCTATTTGACTTTTCTACTCACAAGATGGGGGGA 1800
Qy 1801 ACATGCAAGTGTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAAAATATC 1860
Db 1801 ACATGCAAGTGTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAAAATATC 1860
Qy 1861 TCAGAAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 1861 TCAGACAAATTTGACCAAAATTAAGAGGACGACAAAGAGGGGACTGGTTGGGGCTG 1920
Qy 1921 GGTGGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGATCTCTCTACTA 1980
Db 1921 GGTGGTAAATGGTGACATCTGACTGGGGTGTCTTCTTAACTTTGGGCAATTTTCTCTACTA 1980
Qy 1981 TTATCTATAGTGTCTGATTTGCTCTGCTGCTGATCTGCTATCTTCACTTAAATACATT 2040
Db 1981 TTATCCATAGTGTCTGATTTGCTCTATCTCTATCTTATTTGCTATCTTAAATATATC 2040
Qy 2041 GGATGA 2046
Db 2041 GGATGA 2046
RESULT 14
ABT13451
ID ABT13451 standard; DNA; 1841 BP.
XX
AC ABT13451;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein MBGV-GP1/EBOV-GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
XX WPI; 2003-040651/03.
DR P-PSDB; ABJ18473.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX

CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX

SQ Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;

Query Match 48.7%; Score 995.6; DB 10; Length 6902;

Best Local Similarity 69.5%; Pred. No. 2e-275;

Matches 1352; Conservative 0; Mismatches 594; Indels 0; Gaps 0;

Qy	1	ATGAAGACCATGTTCTCTATCATGCTTATCTTAATTCAGGACAAAAATCTCCC	60
Db	1923	ATGAAGACCATGCTGTTTCATCAGCTGTATCTCTGATCCAGGGGATCAAGACCTGCCC	1982
Qy	61	ATTTTATGAGATGAGTGTAGTAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAACT	120
Db	1983	ATCTGGAGATCGCCAGCAACACAGCCCGCAGACGTGGACAGGTGTGACGGGCACC	2042
Qy	121	CTCAGAGACAGAAAGACGTCTCATCTGATGGGATTTCACTGAGTGGGCAAAAAGTTGCT	180
Db	2043	CTGAGAGAGACGAGGAGGTGCACCTGATGGGCTTTCACCTGAGCGGCGCAGAAAGTGGCC	2102
Qy	181	GATTCCTCTTGGGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
Db	2103	GACAGCCCTCTGGAGGCGCAGAAAGGTGGGCTTTTCAGGACGCGGCTGCCCCCAAGAAC	2162
Qy	241	GTGTAGTACACAGAGGGGAGGAGCCAAACATCTACATATATAGTGTAAACGATCCC	300
Db	2163	GTGGAGTACACCGAGGCGGAGGCGCAAGCTGTCTACACATCAGGTGACGCCCC	2222
Qy	301	TCTGAAAAATCTTCTGTATGATCCTCTCTACCAATCCGTGATCTCCGAAATGCAAA	360
Db	2223	AGCGCAAGAGCTGCTGTGGACCTCCCAACATCAGGACTACCTTAAGTGCAAG	2282
Qy	361	ACTATCCATATATTCAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG	420
Db	2283	ACCATCCACCATCCAGGCGCAGAACCTTCACGCCAGGGGATCGCCCTGCACCTGTGG	2342
Qy	421	GGAGCATTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC	480
Db	2343	GGCGCTTCTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC	2402
Qy	481	ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG	540
Db	2403	ACCGAGGGCAACATCGCGCCATGATCGTTTACAAAGACCGTGCAAGATGATCTTCAGC	2462
Qy	541	CGGCAAGGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATTTGGAAGT	600
Db	2463	AGGCAAGGCGCAGGCTACAGGCACATGAACCTTGACCAAGACCAACATCTCTGGAC	2522
Qy	601	AGTAAACGGAACCAAAACGAAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAA	660
Db	2523	ACAACGCGCACCCAGACCAACGACACCGGCTGCTTGGGCGCCCTGACAGGTATCAAC	2582
Qy	661	ACAAAGAACCAAAATGTGCTCCGTCCTCAAAATACCTTCCACCACTGCCACAGCCGCTCG	720
Db	2583	ACCAAGAACCAAGACCTGCGCCCCCAGCAAGATCCCGAGCCCTGCCCCACCGCCAGG	2642
Qy	721	GAGATCAAACTCACAGACCCCAACTGATGCCACCAAACTCAATACCAAGCCCAAGC	780
Db	2643	GAGATCAAGCCCAACAGACCCCAACTGATGCCACCAAACTCAATACCAAGCCCAAGC	2702
Qy	781	AGTGTATGATGAGGACTCTGCAACATCCGGCTCAGGGTCCGAGAAACGAGAACCCACACA	840
Db	2703	AACGACGACGAGGACTGTATACACGCGGACGCGGACGCGGCGGACGAGGACCTTACAC	2762

Qy	841	ACTTCTGATGCGGTCCACCAAGCAGGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA	900
Db	2763	ACCAAGCAGCGCGTGACCAAGCAGGGGCTTGAGCAGCACCATGCTCTTACCCCTAGCCCT	2822
Qy	901	CAACCAAGCAGCGCACAGCAGAGGAGAAACAAACAAACCAATTCCTCAAGATGTGTGACT	960
Db	2823	CAGCCAGCACCTCTAGCAGGAGGGCAACACCGACCAAGCAGCAGGCGGACCGGTGACC	2882
Qy	961	GAACTAGACAAAAATAACACAACTGCAACACCGTCCATGCCCCCTCATPAACACTACACA	1020
Db	2883	GAGCCCAACAGACCAACACCCAGCCAGCCAGCAGTGTCTCTCAACACCAACCCGCC	2942
Qy	1021	ATCTCTACTAACAACCTCCAAACACAACTTCAGCACTCTCTGCACTTACCAAAAC	1080
Db	2943	ATCAGCACCAACCAACACCAAGCAAGCAAACTTCAGCACCTTGAGCGGTGAGCGCTGAGC	3002
Qy	1081	ACCACCAATGACACACACAGACGACCAATCACTGAAATAGCAAAACAGTGCCTCCGCTCG	1140
Db	3003	ACCACCAATGACACACCAAGACACCGCCACCGGAAAGAGCAGACCAAGCGCCCTAGC	3062
Qy	1141	ATAACAAACCTCTCCAAACCGGAAATCCCAACACAGCAAAAGACACAGCAGCAAAAAA	1200
Db	3063	AAGACCACTCTCTCCACCGCAACCTGACCAACCGCAAGAGCACCACCAACACCAAG	3122
Qy	1201	GGCCCGCCACAAACCGGACCAACACGACCAATGAGCAATTTCAACAGTCTCTCCCGCACC	1260
Db	3123	GGCCCGCCACCAACCGGCGCTTAAATGACCAACCGGCACTTGACCGCCCGCAGCCCGCACC	3182
Qy	1261	CCAGCTCGAGTCGACCAACATCTTGTATATTTTCAAGAAAGAGCGATCGATTTCTGGAAA	1320
Db	3183	CCCAACCCCAACCCAGCAGCTGTGTATCTTCAAGAAAGAGGAGGATCTCTGTGGAGG	3242
Qy	1321	GAAGGTGATATATTTCCGTTTTTATAGTGGTTTAAATACTGAAATTTGATTTTGATCCA	1380
Db	3243	GAGGCGGATATGTTTCCCTCTCGAGCGCTGATCAAGCGCCCTATCGACTTCGACCCC	3302
Qy	1381	ATCCAAACACAGAAACAACTCTTTTGTATGATCTCCAGCTTTAATACTTCAACTAATGAG	1440
Db	3303	GTGCGCAACACCAAGACCATCTTCGACGAGACGACGAGCGGCGCCGCGGAGG	3362
Qy	1441	GAAACACACATCCCGGAAATATCAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA	1500
Db	3363	GACGACGACCGCAGCCCAACATCAGCTGACCTGAGCTACTTCCCAACATCAACGAG	3422
Qy	1501	GATCTGCTACTCTGGGGAAAAACGAGAAATGATTTGTATGACAGGTTTGAGGATTTGGAGT	1560
Db	3423	AACACCGCTTACGCGCGGAGACGAGAACGACTGCGACGCGGAGCTGAGGATCTGGAGC	3482
Qy	1561	GTGAGGAGGACGATTTGGCGGAGGGCTTGTAGCTGGATACCAATTTTGGCCCTCGAATC	1620
Db	3483	GTGAGGAGGAGGACCTGGCGCGCGGCTGAGCTGGATTCCTTCTCGCCCGCGCATC	3542
Qy	1621	GAAGGACTATATCTGCGGTTTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG	1680
Db	3543	GAGGCGCTGTACACCGCGCGCTGATCAAGAACCAAGAACACCTGCTGTGCAAGCTGAGG	3602
Qy	1681	CGCTTAGCTAATCAAACTGCTTAAATCTTGGAGCTCTTGTAAAGGGTCAACACCGAGAA	1740
Db	3603	AGGCTGGCCCAACGACCGCCAGAGCCTGAGCTGCTGTGAGGCTGACCAACGAGGAG	3662
Qy	1741	AGGACATTTTCTTAATCAATAGGATGCAATTTGATTTTTTGTCTTTCAGAGTGGGGCGGA	1800
Db	3663	AGGACCTTTCAGCTGTATCAACAGGACCGCATCGACTTCTGCTGACCAAGGTTGGGGCGGC	3722
Qy	1801	ACATGACAGGTGCTAGGACCTGATTTGTGATAGGAATAGAGATCTATCTATAAAATATC	1860
Db	3723	ACCTGCAAGGTGCTGGGCGCCGACTGCTGCAATCGGATCGAGGACCTGAGCGAGAACATC	3782
Qy	1861	TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA	1920
Db	3783	AGCGAGCAGATCGACCATCAAGAGGACGAGGAGGAGGAGGACCGCGCTGGGGCGCTG	3842
Qy	1921	GGTGGCAAAATGGTGACATCTGACTG	1946

Db 3843 GCGGCAAGTGTGGACCAGCGACTG 3868
|| ||||| ||||| |||||

Search completed: July 31, 2005, 16:57:23
Job time : 1078.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6679.18 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-5

Perfect score: 2046

Sequence: 1 atgaagaccacatgttcct.....tcactaaatcattggtgatga 2046

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	99	4.8	895	9	CNS0071A
C 2	93.2	4.6	1101	9	CNS00FXE
C 3	87.2	4.3	922	9	CNS0073W
C 4	83.4	4.1	884	9	CNS006U0
C 5	80.8	3.9	1101	9	CNS00LO0
C 6	80.6	3.9	1101	9	CNS00LT2
C 7	79.2	3.9	909	9	CNS00JTL
C 8	77.8	3.8	1225	9	CNS0166K
C 9	75	3.7	837	9	AG136151
C 10	73.6	3.6	1965	9	CL090566
C 11	72.6	3.5	939	9	CNS006U0
C 12	72.2	3.5	1787	9	CG754239
C 13	71.8	3.5	1626	9	AG131986
C 14	71.4	3.5	1455	9	AG382036
C 15	70	3.4	1101	9	CNS017YH
C 16	69.2	3.4	902	9	CNS006QP
C 17	68.4	3.3	1101	9	CNS017ZT
C 18	68.2	3.3	861	9	CNS0075A
C 19	68	3.3	939	9	CNS00CNG
C 20	67.4	3.3	1559	9	CG756460
C 21	67	3.3	796	8	B12592
C 22	67	3.3	993	9	CNS005N6
C 23	67	3.3	1319	9	AG371230
C 24	67	3.3	1667	9	CG756610

C	25	66.2	3.2	470	4	BG786319	SEAMC006
	26	66	3.2	723	8	AZ627893	1M0469118
	27	66	3.2	1025	9	CL487193	SAIL_449
	28	65.6	3.2	527	9	CE480293	t1gr-ges-
	29	65.6	3.2	860	9	CNS018FL	Drosophil
	30	65.4	3.2	913	6	CA986290	AGENCOURT
	31	65.4	3.2	925	7	CK425533	AUF IpTee
	32	65	3.2	769	8	BZ579780	mh2_769
	33	65	3.2	869	9	AG137111	Pan trogl
	34	65	3.2	1147	8	CC187235	CH261-152
	35	64.8	3.2	513	9	CNS02CMF	AL191328
	36	64.8	3.2	633	7	CV070818	CS_gil_34
	37	64.6	3.2	937	9	CNS006ST	AL065880
	38	64.6	3.2	1153	9	CNS07BVP	AL438395
	39	64.6	3.2	1345	8	AQ743328	T3 end of
	40	64.6	3.2	1613	8	BZ557155	BZ557155
	41	64.4	3.1	917	9	CL466449	SAIL_1256
	42	64.2	3.1	1131	9	CNS03AWV	AL235768
	43	64	3.1	2263	3	CR698468	Tetraodon
	44	63.8	3.1	732	9	AG135778	Pan trogl
	45	63.8	3.1	1183	9	AG136828	Pan trogl

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL066286

VERSION
AL066286.1 GI:4945153

KEYWORDS
GSS.

SOURCE
Drosophila melanogaster (fruit fly)

ORGANISM
Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 895)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammose in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1. 895

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR14B09"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match 4.8%; Score 99; DB 9; Length 895;

Best Local Similarity 24.1%; Pred. No. 2.8e-16;

[illegible]

RESULT 2	CNS00FXE	1101 bp	DNA	linear	GSS 03-JUN-1999
	CNS00FXE/c				
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence TtT3 end of BAC: BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				

ACCESSION	AL071370	
VERSION	AL071370.1	GI:4951210
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
	Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 1101)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :	
	BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr	
	- Web : www.genoscope.cns.fr)	
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila/bac.htm .	

FEATURES	source
found at	location/Qualifiers
	1. .ilol1
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACR32C19"
	/clone_lib="RPCI-98"

[illegible]

RESULT 3

CNS0073W/c	CNS0073W	922 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster	genome survey			
DEFINITION	BAC14D09 of RPCI-98 library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.				
ACCESSION	AL066784				
VERSION	AL066784.1	GI:4945247			
KEYWORDS	GSS.				
SOURCE	<i>Drosophila melanogaster</i> (fruit fly)				
ORGANISM	<i>Drosophila melanogaster</i>				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; <i>Drosophila</i> .				
REFERENCE	1 (bases 1 to 922)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)				

COMMENT

new : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Ooegawa and Aaron Mammoss in Piter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
Location/Qualifiers
1. .922
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 4.3%; Score 87.2; DB 9; Length 922;
Best Local Similarity 22.1%; Pred. No. 6.5e-13;
Matches 78; Conservative 145; Mismatches 130; Indels 0; Gaps 0;
Qy 930 CAACACAAACCATTCCTGAGTGTGCTGAGTAACTAGACAAATAATACACAACTGACACA 989
Db 905 MMMCM 846
Qy 990 ACCTGTCATGCCCCCTATACACTACCAATCTCTACTAACCAACCTCCAAACACAA 1049
Db 845 MCMCMACM 786
Qy 1050 CTTGAGCACTCTCTGTCACCATTTACAAACACCAACCAATGACAAACACAGACACAAT 1109
Db 785 MMMCM 726
Qy 1110 CACTGAAATGAGCAACAGCTGCCCCCTCGATACCAACCTGCTCCAAACGGAAATCC 1169
Db 725 MMMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMA 666
Qy 1170 CACCACGAAAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1229
Db 665 AAMAMMCCCCAACACM 606
Qy 1230 AAATGAGCATTTCCAGCTCTCTCCCCCACCCTGCTGCTGCTGCTGCTGCTGCTGCT 1282
Db 605 MMACM 553

RESULT 4

CNS006U0/C
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION
AL065923.1 GI:4944891

KEYWORDS

SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org>. The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
Location/Qualifiers
1. .884
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 4.1%; Score 83.4; DB 9; Length 884;
Best Local Similarity 20.1%; Pred. No. 7.7e-12;
Matches 78; Conservative 171; Mismatches 138; Indels 1; Gaps 1;
Qy 855 CACCAACCAAGGGTTTCATCAACAATGTCACCCACTCCCTCACCACAAACCAAGCAGCC 914
Db 881 MMMVMMHHTKKKKTTTHMMVMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 822
Qy 915 ACAGCAAGGAGGAAACAAACCAATCCCAAGATGCTGTGACTGAATAGACAAAAA 974
Db 821 MMM 762
Qy 975 TAACACAACTGACACACCGTCCATGCCCCCTCATACACTTACCAATCTCTACTACAA 1034
Db 761 MCMMMMMMMMMMMMMMMMMMMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 702
Qy 1035 CACCTCAAAACACAACCTTACGACTCTCTGTGACCACTTACAAACACCAACCAATGACAA 1094
Db 701 MMAM-CHCM 643
Qy 1095 CACACAGACGACATCTACTGAAATGAGCAAAACAGTGCCCTCGATATAACAACTGCCC 1154
Db 642 CCCCCCMMAACAAACACACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 583
Qy 1155 TCCAAACGGGAAATCCACACGACGACGACGACGACGACGACGACGACGACGACGAC 1214
Db 582 CAMCTCACCACAAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAH 523
Qy 1215 GGACACAAACACGACAAATGAGCATTTTC 1242
Db 522 KCBCMYTBCCCCCCMMMAAMMAAKCKC 495

RESULT 5

CNS00LO0

LOCUS

DEFINITION

Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION
AL068607

KEYWORDS

SOURCE
GI:4958689

ORGANISM

Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr

COMMENT

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org>. The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of


```
Db 1048 ACMCMCMCMCCACAAAMAMCAATATAAAMGACACACMCMMAAAMACACCCCMCMCCACAAAA 989
Qy 907 AGCAGCCACAGCAGGAGGAAACACACAAACCAATCCCAAGATGTGTGACTGAACATA 966
Db 988 CCCACCCCAACACACACACACACACACACACACACACACACACACACACACACACACAC 929
Qy 967 GACAAAA---ATAACACAACTGCAACACCGTTCATGCGCCCTCATTAACACTACACAAAT 1022
Db 928 CMMAACCCCCACACACACACACACACACACACACACACACACACACACACACACACAC 869
Qy 1023 CTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTGACCACTTACAAAAACAC 1082
Db 868 CMMAACCMMAAAWCAAAACACACACACACACACACACACACACACACACACACACAC 809
Qy 1083 CACCAATGACACACACAGACACACATCACTGAAATGAGCAAAACAGTGGCCCCCTCGAT 1142
Db 808 CMCCCAACCMMAAAWCAAAACACACACACACACACACACACACACACACACACACAC 749
Qy 1143 AACAAACCTGCTCCAAACGGGAAATCCCAACACACAGCAAGAGCAGCAGCAAAAAAGG 1202
Db 748 MAATAAAKYMAAAACAAKAAWAAWYCMCCGRARVMSSSSGVSVVSVVVVVVAGAA 689
Qy 1203 CCCCACACACAGCACCACACACACACACAAATGATTCACCACTCTCCGCC 1257
Db 688 MAATAAAARAGGCGCAAMMCCCCSCMAAAASCSMCCRGSGSVSSCCCCCS 634

RESULT 9
AG136151
LOCUS Pan troglodytes DNA, clone: PTB-149K08.F, genomic survey sequence.
ACCESSION AG136151
VERSION AG136151.1 GI:16665829
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
TITLE Totoki, Y., Watanabe, H. and Sakaki, Y.
JOURNAL BAC end sequences of Library PTB
AUTHORS Unpublished
2 (bases 1 to 837)
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
JOURNAL Totoki, Y., Watanabe, H. and Sakaki, Y.
AUTHORS Direct Submission
TITLE Submitted (02-AUG-2001) Asso Fujiyama, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
AUTHORS 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .837
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-149K08.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
3.7%; Score 75; DB 9; Length 837;

ORIGIN
Query Match
```

```
Best Local Similarity 47.1%; Pred. No. 1.8e-09;
Matches 264; Conservative 0; Mismatches 292; Indels 5; Gaps 1;
Qy 727 AAATCTCAACAGCAGCCCAACTGATGTCACCACTCAATACCGGACCCAGCAGTGTAT 786
Db 261 AAACCAACCAACCAACCAACCAACCAACCACTTATCCACCCACCCACCCACA 320
Qy 787 GATGAGGACCTCGCAACATCCGGTCTCAGGGTCCGGAGAACGAGAACCCCAACAATTCT 846
Db 321 AACACACACACAAAAACACAAACCAACCAACCAACCAACCAACCAACCAACCAAC 380
Qy 847 GATGGGTCTACCAAGCAGGCGCTTTCATCACAATGCCACCACTCTCTCCACCAACA 906
Db 381 AACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 440
Qy 907 AGCAGCCACAGCAGGAGGAAACAAACAAACCAATTCCTCCCAAGATGTGTGACTGAAC 966
Db 441 AGCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 500
Qy 967 GA-----CAAAATAACACAACTGCACACCGTCCATGCGCCCTCATTAACACTACCA 1021
Db 501 CAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 560
Qy 1022 TCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGACCACTTACAAAAA 1081
Db 561 CCACAAACCAACATCCCCAAAACCAACACACACCAACCAACCAACCAACCAACCA 620
Qy 1082 CCACCAATGACACACACAGCAGCAATCACTGAAATGAGCAAAACAGTGCCCCCTCGA 1141
Db 621 CCACCAATGACACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 680
Qy 1142 TAACACACCTGCTCCAAACGGGAAATCCCAACAGCAAGAGCAGCAGCAAAAAAG 1201
Db 681 AACAACTCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 740
Qy 1202 GCCCGGCAACACGGCACCACCAACAGCAATGAGCAATTCACCACTCTCTCCCAAC 1261
Db 741 ACCAAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 800
Qy 1262 CCAGCTGAGTGACCAACATC 1282
Db 801 AAACCAAAACCAACCAACCC 821

RESULT 10
CL090566/c
LOCUS ISB1-17017_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-17017,
DEFINITION genomic survey sequence.
ACCESSION CL090566
VERSION CL090566.1 GI:40584201
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1965)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 781
High quality sequence stop: 832.
Location/Qualifiers
1. .1965
source
```

```

/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8164"
/clone="ISB1-17017"
/clone_lib="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

```

ORIGIN

```

Query Match      3.6%; Score 73.6; DB 9; Length 1965;
Best Local Similarity 48.8%; Pred. No. 6e-09;
Matches 199; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 875 CAACATGCCACCCACTCCCTCACCACCAACCAGCAGCCGACAGGAGGAACAACA 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 CACCAACACCAACACACCCACCCACCAACACCCACCCACCAACCCACACACCC 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 935 CAACCAATTCCTCCCAAGATGCTGTGACTGAATAGACAAAATAACCAACTGCACAAACCGT 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 ACACCAACACCAACCCACCCCAACCAACCAACCAACCAACCAACCAACCAACAAAC 1066
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 CCATGCCCTCTATACACTACCACTCTCTACTTAACAACACACTCTTCCAAACACAACTTCA 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1065 CCCCCACACACCCACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1055 GCATCTCTCTGCACATTACAAAACACCAACCAATGACACACAGAGACCAATCATCTG 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1005 CCCCACAAACACACACCCACCCACCCACCAACCAACCAACCAACCAACCAACCAACCA 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1115 AAAATGACCAACACAGTGCCTCGTATGAACCAACCTGCTTCCACGGGAATCCACCA 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 CACACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1175 CAGCAAGAGCAGCAGCAGCAAAAGCCCGCCACCAACCGCAGCAACCAACGACAAATG 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 CAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1235 AGCATTTACAGTCTCTCCCGCCACCCGAGTGCACTGCACAAATC 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 AAAAAAAAACACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11

```

CNS006UU/c      939 bp DNA linear GSS 03-JUN-1999
LOCUS           Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION      BAC14L09 of RPCI-98 library from Drosophila melanogaster (fruit
SOURCE          fly), genomic survey sequence.
ACCESSION       AL065953.1 GI:494921
VERSION         GSS.
KEYWORDS        Drosophila melanogaster (fruit fly)
SOURCE          Drosophila melanogaster
ORGANISM        Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE       1 (bases 1 to 939)
AUTHORS         Genoscope.
TITLE           Direct Submission
JOURNAL         Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT         Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org The BDGP Drosophila
                melanogaster BAC library was prepared by Kazutoyo Osoegawa and
                Aaron Mammosier in Pieter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's

```

Pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

```

source
1..939
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BAC14L09"
    /clone_lib="RPCI-98"
    /note="end : T7"

```

ORIGIN

```

Query Match      3.5%; Score 72.6; DB 9; Length 939;
Best Local Similarity 33.8%; Pred. No. 9.2e-09;
Matches 124; Conservative 92; Mismatches 150; Indels 1; Gaps 1;

QY 795 CCTCGCAACATCCGGCTCAGGGTCGGAGACGAGAGAGACGACGMDGMSRVCGGASVGVCA 854
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 865 MSSCMSACCCACRCMGMCAGCGCCCRAGACRAGAGAGACGACGMDGMSRVCGGASVGVCA 806
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 855 CACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCACCAACCAAGCACGCC 914
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 805 VGSVVMACGAGSAGVGVGRCACACGVMVVCASCSMACMAMAMRMCAMAVASAAANAA 746
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 915 ACAGCAAGGAGGAACAACAACCAATTCCTCCAGATGCTGTGACTGAACTAGACAAAAA 974
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 745 VAHVAGAVARGGGAAMGCRGAACRMAACMWCACAMAMARMSACGSAACAMMAACAS 686
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 975 TAACAACTGCACACCGTCCATGCCCTCCATTAACACTACCACTACCAATCTCTACTAACAA 1034
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 685 VARMACAMAMACAMMAAAACSMRCVMCVSRACAMACCAACCAACACACACCCCMCMCA 626
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1035 CACCTTCCAAACCAACAATTCA-GCACTCTCTCTGCACCACTTACAAAACCAACCAATGACA 1093
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 625 CCCCMCMCMMAACCAACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 566
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1094 ACACACAGAGCAACATCTCTGAAATAGCAAAACAGTGGCCCCCTCTGATAACACCCCTGC 1153
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 565 CMACMCCCTATCMMAAACACAAAMAMAMACMAAACMAACMAACMAACMAACMAACMAAC 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1154 CTCCAAC 1160
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 505 MCCCCRC 499
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

```

CG754239/c      1787 bp DNA linear GSS 24-OCT-2003
LOCUS           P049-3-D12.ya Ppa EcoRI BAC library Pristionchus pacificus genomic,
DEFINITION      genomic survey sequence.
ACCESSION       CG754239
VERSION         CG754239.1 GI:37979530
KEYWORDS        GSS.
SOURCE          Pristionchus pacificus
ORGANISM        Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
                Neodiplogasteridae; Pristionchus.
REFERENCE       1 (bases 1 to 1787)
AUTHORS         Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
                Buntjer,J., van der Meulen,M. and Sommer,R.J.
                An integrated physical and genetic map of the nematode Pristionchus
                pacificus
                Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL         Evolutionary Biology
MEDLINE         Max-Planck-Institute for Developmental Biology
                Spemannstr. 37-39, Tuebingen D-72076, Germany
                Tel: 00497071601371
                Fax: 00497071601498
COMMENT         12884007
                PUBMED

```

Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. 1787
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match 3.5%; Score 72.2; DB 9; Length 1787;
Best Local Similarity 48.9%; Pred. No. 1.5e-08;
Matches 194; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

QY 883 CCACCACTCCCTACCAACCAAGCAGCGCCACAGGAGGAGGAAACACACAAACCAT 942
|||||
Db 758 CCACCAACCATATCACACACACACACACCTCCCATCAACACAAACACACACACCA 699
|||||
QY 943 TCCCAGATGCTGTGACTGACTAGCAAAATACACAACTGCACACCGTCCATGCC 1002
|||||
Db 698 ACACCAAGAGACACTTAACCAACACACACACCCGCCCAAAAAAAGCCACCC 639
|||||
QY 1003 CCTATAACACTATCACCAATCTCTACTAACACACTCTCAACACAACTTCAGCACTCTC 1062
|||||
Db 638 AACACCCGCCACACACACCAAAACAAACCAACACACACATACACACACACAC 579
|||||
QY 1063 TGTGACCATTTACAAACACCAACCAATGACACACACAGAGACAAATCACTGAAATGAG 1122
|||||
Db 578 TCAACACATCAACAACATACACACACACGCCCATACAACTTCCACCCACACCC 519
|||||
QY 1123 CAAACAGTGGCCCTCGATACACACCTGCTCCAAAGGGAATCCACACACACAAAG 1182
|||||
Db 518 AACACACAAACAAACACACACACACCCCAACCCCAACACACACACACATCC 459
|||||
QY 1183 AGCACAGCAGCAAAAAAGGCCCGCCCAACAGGACCAACACGACCAAAATGAGCATTT 1242
|||||
Db 458 ACCACATCCACACACACACACACACCCGACCAAAACCAACCAACACACACCA 399
|||||
QY 1243 ACCAGTCTCCCCACACCCCGAGCTCGACTGCACAA 1279
|||||
Db 398 ACAACCCACCCACACACACACACACACACACACAC 362
|||||

RESULT 13

AG131986
LOCUS Pan troglodytes DNA, clone: PTB-144D13.R, linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-144D13.R, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. 1626

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-144D13.R"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 3.5%; Score 71.8; DB 9; Length 1626;
Best Local Similarity 39.7%; Pred. No. 1.8e-08;
Matches 238; Conservative 0; Mismatches 362; Indels 0; Gaps 0;

QY 647 AAGAATACAAATCTACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACCATG 706
|||||
Db 156 AAANNAANANNCNCANAAAAACNANAAACNANNNNNNNNNNNNNNNNNNNNNNN 215
|||||
QY 707 CCACAGCCCGTCCGAGATCAAACTCAACAGCACCCCAACTGATGCGACCACTCAATA 766
|||||
Db 216 CCACANNAACNNNAACNANNAACCCNCCACCCNCCNNAACCAACCAAAAAA 275
|||||
QY 767 CCACGAGCCCAAGCAGTGTATGAGGACCTCGCAACATCGGGTCCAGGTCGAGAAC 826
|||||
Db 276 NANCNCCCAAAAACAAAAAANNCNCAAAAAAANNNNCAAAAAAANNNNN 335
|||||
QY 827 GAGAACCCCAACAACTTCTGATGGGTACCAAGCAAGGGGTTTCATCAACATGCGCAC 886
|||||
Db 336 NNNNNCCCGCCNN 395
|||||
QY 887 CCACCTCCCTCACCAACCAAGCAGCCACAGCAGGAGGAGGAAACCAACCAACCATTC 946
|||||
Db 396 CCCAAACCCCAACACCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 455
|||||
QY 947 AAGATGTGTGACTGAACTAGACAAAAATACACAACTGCACAAACCGTCCATGCCCTC 1006
|||||
Db 456 CACANACAAAAACAAAAANCCCAACCAACAAAAAACAACCCCGCCCGCCACACAA 515
|||||
QY 1007 ATAACACTACCAATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTG 1066
|||||
Db 516 NACACACACCCACNANNAACCAAAAAACAAACCCCAACANNAACANWACACCC 575
|||||
QY 1067 CACCATTAACAAACACACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAA 1126
|||||
Db 576 CACAAACCCCAACACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 635
|||||
QY 1127 CCAGTGGCCCTCGATACAAACCTCGCTCCCAAGGGAAATCCCAACAGCAAGACA 1186
|||||
Db 636 CCCCAACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 695
|||||
QY 1187 CCAGCAGAAAAAAGGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1246
|||||
Db 696 ACCCNCAACCAACCAACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 755
|||||

RESULT 14

AG382036/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AG382036 1455 bp DNA linear GSS 03-JUN-2004
Mus musculus molossinus DNA, clone: MSMg01-192F03.TJ, genomic survey
sequence.
AG382036
AG382036.1 GI:47993241
GSS.
Mus musculus molossinus

ORGANISM Mus musculus molossinus
Bumkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1

REFERENCE Hattori.M., Toyoda.A., Noguchi.H., Kojima.T. and Sakaki.Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL 2 (bases 1 to 1455)
REFERENCE Hattori.M., Toyoda.A., Noguchi.H., Kojima.T. and Sakaki.Y.
AUTHORS Direct Submission
TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
JOURNAL and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel:81-45-503-9111, Fax:81-45-503-9170
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Taikuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY Vector : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES Location/Qualifiers
source 1..1455
 /organism="Mus musculus molossinus"
 /mol_type="genomic DNA"
 /sub_species="molossinus"
 /db_xref="taxon:57486"
 /clone="MSMg01-192F03.TJ"
 /sex="male"
 /tissue_type="mixture of kidney and spleen"
 /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 3.5%; Score 71.4; DB 9; Length 1455;
Best Local Similarity 48.8; Pred. No. 2.3e-08;
Matches 222; Conservative 0; Mismatches 231; Indels 2; Gaps 1;

Qy 830 AACCCACACAATCTTGATCGGTACCAGCAAGGGCTTCATCAACTAATGCCACCCA 889
 ||||| ||| |||
Db 1414 AACACACACACACACCCACCCACCGGACACCCCAACCCACACACACACACCAACA 1355

Qy 890 CTCCTTCCACCAACCAAGCAGCCACAGCAAGGAGGAACAAACAAAACCATTTCCAAAG 949
 ||||| ||| |||
Db 1354 GACCCACCCACACACACACAAACCCCAACACACACACACACACACAAACAAAAACCAA 1295

Qy 950 ATGCTGTGACTGACTAGACAAAATAACAACTGCACACCTTCATGCCCTCAT 1009
 ||||| ||| |||
Db 1294 CACAAAAACACACACCAACCAACCAAAAAAACAGAACCAACACCCCACCCACCAAC 1235

Qy 1010 ACATTACACAATCTCTACTTAACAAACACTTCCAAACACAACTTCAGCACCTCTCTGCGAC 1069
 ||||| ||| |||
Db 1234 ACAACACAGCACAACACACACACCCCAACAGCCCAACCAACCCCGGACCCCAACACACAC 1175

Qy 1070 CATTTAAAAACACCAACCAATGACAAAC--ACACAGAGCACATATCTGAAATATGACAAAC 1127
 ||||| ||| |||
Db 1174 CAACACACACCCACCAACCAACCAACCAACCAAAAAACAAAGCCCAACAAACCCCAACAAA 1115

Qy 1128 CAGTGCCCCCTCGATACAAACCTTGCTTCCACGGGAAATCCACACACAGCAAGAGCAC 1187
 ||||| ||| |||
Db 1114 CACAACAACAAACACACACACACACCCCAACACACCCCAACCAACCAACCAACCAACAA 1055

Qy 1188 CAGCAGCAAAAAAGCCCGCCACACCGGCACAAACGCAACACGACAAATGAGCATTTCCACAG 1247
 ||||| ||| |||
Db 1054 CACCACCAACCAACCAACCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAAC 995

Qy 1248 TCCTCCCCCCCCCCCCAGCTCGACTGCACAAACATC 1282

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 327.71 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacatgttctct.....tcactaaatacatggatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1841.2	90.0	2164	3	US-08-760-615-3
2	1841.2	90.0	11460	4	US-09-336-910A-1
3	1510	73.8	2247	3	US-08-760-615-5
4	131.2	6.4	2172	3	US-08-760-615-1
5	131.2	6.4	2298	4	US-09-650-086A-1
C 6	74.6	3.6	152132	4	US-09-949-016-13845
C 7	74.6	3.6	152145	4	US-09-949-016-12371
8	60.4	3.0	34230	4	US-09-949-016-12052
9	60.4	3.0	128470	4	US-09-949-016-13765
10	57.6	2.8	5511	3	US-08-928-361B-2
11	57.6	2.8	5511	4	US-09-588-995A-2
12	57.6	2.8	7334	3	US-08-928-361B-1
13	57.6	2.8	7334	4	US-09-588-995A-1
C 14	54.2	2.6	30656	4	US-09-949-016-14613
C 15	54	2.6	522	4	US-09-949-016-103758
16	53.6	2.6	5163	3	US-08-700-651-1
17	53.6	2.6	5163	3	US-08-928-361B-4
18	53.6	2.6	5163	4	US-09-588-995A-4
19	53.6	2.6	5318	3	US-08-700-651-2
20	53.6	2.6	5318	3	US-08-928-361B-3
21	53.6	2.6	5318	4	US-09-588-995A-3
C 22	52.6	2.6	271134	4	US-09-949-016-12705
C 23	52.6	2.6	305491	4	US-09-949-016-17550
C 24	52.4	2.6	39154	4	US-09-949-016-12384
C 25	52.4	2.6	39154	4	US-09-949-016-12801
C 26	52.4	2.6	39443	4	US-09-949-016-14326
C 27	52.4	2.6	39443	4	US-09-949-016-14327

28	51.6	2.5	114793	4	US-10-148-806-3	Sequence 3, Appli
C 29	51	2.5	57280	4	US-09-949-016-11796	Sequence 11796, A
C 30	51	2.5	57280	4	US-09-949-016-12843	Sequence 12843, A
C 31	51	2.5	57280	4	US-09-949-016-12844	Sequence 12844, A
C 32	51	2.5	57280	4	US-09-949-016-12846	Sequence 12846, A
C 33	51	2.5	57280	4	US-09-949-016-13542	Sequence 13542, A
C 34	51	2.5	57280	4	US-09-949-016-13543	Sequence 13543, A
C 35	51	2.5	57280	4	US-09-949-016-13544	Sequence 13544, A
C 36	51	2.5	57280	4	US-09-949-016-13545	Sequence 13545, A
C 37	51	2.5	57280	4	US-09-949-016-14633	Sequence 14633, A
C 38	51	2.5	57280	4	US-09-949-016-14634	Sequence 14634, A
C 39	51	2.5	57280	4	US-09-949-016-14635	Sequence 14635, A
C 40	51	2.5	57280	4	US-09-949-016-14636	Sequence 14636, A
C 41	51	2.5	57280	4	US-09-949-016-14637	Sequence 14637, A
C 42	51	2.5	57280	4	US-09-949-016-14638	Sequence 14638, A
C 43	51	2.5	57280	4	US-09-949-016-14639	Sequence 14639, A
C 44	51	2.5	57280	4	US-09-949-016-14640	Sequence 14640, A
C 45	48.8	2.4	60990	4	US-09-949-016-14080	Sequence 14080, A

ALIGNMENTS

RESULT 1
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760.615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3

Query Match Best Local Similarity 90.0%; Score 1841.2; DB 3; Length 2164; Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;									
QY	1	ATGAAGACCATGTTTCCCTTATAGTCTTATCTTAAATTCAGGGACAAAAATCTCC	60	QY	1081	ACCACCAATGACAAACACACAGAGCACAATCACTGAAATGAGCAAAACAGTCCCTCG	1140		
DB	119	ATGAAGACCATGTTTCCCTTATAGTCTTATCTTAAATTCAGGGACAAAAATCTCC	178	DB	1199	ACCACCAATGACAAACACACAGAGCACAATCACTGAAATGAGCAAAACAGTCCCTCG	1258		
QY	61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGTTCGGAAT	120	QY	1141	ATAACAACCCCTCCAAACGGGAAATCCCAACAGCAAAAGAGCACCAGCAGCAAAA	1200		
DB	179	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGTTCGGAAT	238	DB	1259	ATAACAACCCCTCCAAACGGGAAATCCCAACAGCAAAAGAGCACCAGCAGCAAAA	1318		
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGATTCACACTGAGTGGGCAAAAGTTGCT	180	QY	1201	GGCCCGGCCACAAAGCGCACCAAAACAGCAAAATGAGCAATTCACCACTCTCCCCCAACC	1260		
DB	239	CTCCAGAGACAGAGAGAGCTCCATCTGATGGATTCACACTGAGTGGGCAAAAGTTGCT	298	DB	1319	GGCCCGGCCACAAAGCGCACCAAAACAGCAAAATGAGCAATTCACCACTCTCCCCCAACC	1378		
QY	181	GATTCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT	240	QY	1261	CCCAGCTCGACTGCAACAACATCTTTGTATATTTTCAGAGAAAGCGATCGATTTCTCGAAA	1320		
DB	299	GATTCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT	358	DB	1379	CCCAGCTCGACTGCAACAACATCTTTGTATATTTTCAGAGAAAGCGAATATCTCTCGGAGG	1438		
QY	241	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	300	QY	1321	GAAGTGATATATTTCCCGTTTTTAGATGGGTATAATAACTGAAATTCGAATTTTGATCCA	1380		
DB	359	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	418	DB	1439	GAAGGCGACATGTTCCCTTTTCGTGATGGGTATAATACTCTCCAATGATTTTGACCCA	1498		
QY	301	TCCTGAAAAATCCTTCTGTAGATCTCTACCAACATCCGTGACTATCCGAAATCGAAA	360	QY	1381	ATCCCAACACAGAAAAACAATCTTTTGATGAATCTCCAGCTTTTAATATCTTCAACTAATGAG	1440		
DB	419	TCGGAANAATCCTTCTGTGTAGATCTCTACCAACATCCGTGACTATCCGTAATCGAAA	478	DB	1499	GTTCCAATATCAAAAAACAATCTTTTGATGAATCTCTAGTTCCTGGTCCCTCGGCTGAGGAA	1558		
QY	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420	QY	1441	GAAACAACACTCCCCCGAATATCAGTTTAACTTCTCTTATTTTCTGATATAAAATGGA	1500		
DB	479	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	538	DB	1559	GATCAACATGCTCTCCCAATATATAGTTTAACTTTATCTTATTTTCTTAATAATAATGAG	1618		
QY	421	GGAGCATTTTCTGTATGATCGCATGCTCCCAACAATGTACCGAGGCAAGTCTTC	480	QY	1501	GATACTGCTACTCTGGGAAAAACGAGAATGATTTGTGATGACAGAGTTGAGGATTTGGAGT	1560		
DB	539	GGAGCATTTTCTGTATGATCGCATGCTCCCAACAATGTACCGAGGCAAGTCTTC	598	DB	1619	AACACTGCTACTCTGAGAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGGAGC	1678		
QY	481	ACTGAAGGGAAACATAGCAGCTATGATTTGTAATAGACAGTGCACAAAATGATTTTCTCG	540	QY	1561	GTGAGAGGACGATTTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTCGGAATC	1620		
DB	599	ACTGAAGGGAAACATAGCAGCTATGATTTGTAATAGACAGTGCACAAAATGATTTTCTCG	658	DB	1679	GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT	1738		
QY	541	CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAATAAATATTCGACAAGT	600	QY	1621	GAAGGACTTATACTGCGGGTTTTAATCAAAAAATCAGAAATTTAGTTTGTAGGTTGAGG	1680		
DB	659	CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAATAAATATTCGACAAGT	718	DB	1739	GAAGGACTTTTACACTGCTGTTTTTAAATTAATAAATCAAAACAAATTTGGTCTGCGAGTTGAGG	1798		
QY	601	AGTAAACGGAAACCAACGAACTGATGATTTTCGGCGCTCTTCAAGAAATCAAAATCT	660	QY	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTAAAGGGTCACAACCCGAGGAA	1740		
DB	719	AGTAAACGGNAACCAACGNAATGACACTGGATGTTTTTCGGCGCTCTTCAAGAAATCAAAATCT	778	DB	1799	CGTCTAGCAATCAAACTGCAAAATCCTTTGGAATCTTATTATGAGAGTCACAACTGAGGAA	1858		
QY	661	ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACACATGCCCCACAGCCGCTCG	720	QY	1741	AGGACATTTTCTTAAATCAATAGCATGCAATTTGACTTTTGTCTTACGAGGTGGGGCGGA	1800		
DB	779	ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACACATGCCCCACAGCCGCTCG	838	DB	1859	AGAACATTTCTTAAATCAATAGACATGCTATTGACTTTCTACTCAAGATGGGGAGGA	1918		
QY	721	GAGATCAAACTCAAGACACCCCACTGATGCGCACCAAACTCAATACCAAGGACCCCAAGC	780	QY	1801	ACATGCAAGGTCTAGGACCTGATTTGTGATAGGAATAGGAATCTATCTAAAAATATC	1860		
DB	839	GAGATCAAACTCAAGACACCCCACTGATGCGCACCAAACTCAATACCAAGGACCCCAAGC	898	DB	1919	ACATGCAAGGTCTTGGACCTGATTTGTGATCGGGATAGAGACTTGTCCAAAAATAT	1978		
QY	781	AGTGTATGAGGACCTCGCAACATCCGGGCTCAGGGGTCGGAGAACGAGAACCCCAACA	840	QY	1861	TCAGAACAAATTCGACAAAAATCAGAAAGGATGAAACAAAAGGAGGAAACTGGCTGGGGTCTA	1920		
DB	899	AGTGTATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACA	958	DB	1979	TCAGAGCAAAATTGACCAAAATTAATAAGAGCGAAACAAAAGAGGGGACTGGTTGGGGTCTG	2038		
QY	841	ACTTCTGATGCGGTCAACGAAGAGGGCTTTTCATCAACAATGCCACCACTCCCTCAGCA	900	QY	1921	GGTGGCAATGGTGGACATCTGACTGGGGTGTCTTCAACAAATTTGGGCACTCTGTCTACTA	1980		
DB	959	ACTTCTGATGCGGTCAACGAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCAGCA	1018	DB	2039	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTTGTCTACTA	2098		
QY	901	CAACCAAGCACGCCACAGCAAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT	960	QY	1981	TTATCTATAGCTGTTTCTGATTTGCTCTGCTGTATCTGTCTGATCTTCTCACTAAATACATT	2040		
DB	1019	CAACCAAGCACGCCACAGCAAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT	1078	DB	2099	TTATCCCATAGCTGCTTGAATGCTCTATCTCTGATTTTGTCTGATCTTTACTATAATATATC	2158		
QY	961	GAACTAGACAAAAATAACAACCTGCAACACCGTCCATGCGCCCTCATATACTACCAACA	1020	QY	2041	GGATGA 2046			
DB	1079	GAACTAGACAAAAATAACAACCTGCAACACCGTCCATGCGCCCTCATATACTACCAACA	1138	DB	2159	GGATAA 2164			

RESULT 2

US-09-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; CURRENT FILING DATE: 1999-06-21
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-09-336-910A-1

Query Match 90.0%; Score 1841.2; DB 4; Length 11460;

Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy	1	ATGAAGACCAATGTTCTTATCAGTCTTATCTTAAATTAAGGACAAAAATCTCCC	60
Db	5940	ATGAAGACCAATGTTCTTATCAGTCTTATCTTAAATTAAGGACAAAAATCTCCC	5999
Qy	61	ATTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCGCGAAT	120
Db	6000	ATTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCGCGAAT	6059
Qy	121	CTCAGAGACAGAGAGCTGATCTGATGGATTCACATGAGTGGGCAAAAGTTGCT	180
Db	6060	CTCAGAGACAGAGAGCTGATCTGATGGATTCACATGAGTGGGCAAAAGTTGCT	6119
Qy	181	GATTCCTCTTGGAGGATCAAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240
Db	6120	GATTCCTCTTGGAGGATCAAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	6179
Qy	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGATCCC	300
Db	6180	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGATCCC	6239
Qy	301	TCTGAAAAATCTTGCTGTTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA	360
Db	6240	TCTGAAAAATCTTGCTGTTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA	6299
Qy	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
Db	6300	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	6359
Qy	421	GGAGCATTTTTCTGTATGATGCAATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	480
Db	6360	GGAGCATTTTTCTGTATGATGCAATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	6419
Qy	481	ACTGAAGGAAACATAGAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCG	540
Db	6420	ACTGAAGGAAACATAGAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCG	6479
Qy	541	CGGCAAGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATTGGACAGT	600
Db	6480	CGGCAAGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATTGGACAGT	6539
Qy	601	AGTAACGGAACGCAAAACGAATGACATCGATGTTTTCGGGCTCTTCAAGAATACAAATCT	660
Db	6540	AGTAACGGAACGCAAAACGAATGACATCGATGTTTTCGGGCTCTTCAAGAATACAAATCT	6599

Qy	561	ACAAAGAACCAACATGCTCCGTCCAAATATACCTCCACACATGCCCCACAGCCGTCGG	720
Db	6600	ACAAAGAACCAACATGCTCCGTCCAAATATACCTCCACACATGCCCCACAGCCGTCGG	6659
Qy	721	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	780
Db	6660	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	6719
Qy	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	840
Db	6720	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	6779
Qy	841	ACTTCTGATGGGTCAACAGCAGGGCTTTCATCAACATGCCACCACTCCCTCACA	900
Db	6780	ACTTCTGATGGGTCAACAGCAGGGCTTTCATCAACATGCCACCACTCCCTCACA	6839
Qy	901	CAACCAAGCAGCCACAGCAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTACT	960
Db	6840	CAACCAAGCAGCCACAGCAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTACT	6899
Qy	961	GAAGTAGACAAAAATAACACAACTGCAACACCGTCCATGCCCCCTCATAACTACACCA	1020
Db	6900	GAAGTAGACAAAAATAACACAACTGCAACACCGTCCATGCCCCCTCATAACTACACCA	6959
Qy	1021	ATCTTACTTAACACACTCCCAACACAACTTTCAGCACTCTCTGACCACTTACAAAC	1080
Db	6960	ATCTTACTTAACACACTCCCAACACAACTTTCAGCACTCTCTGACCACTTACAAAC	7019
Qy	1081	ACCACCAATGACACACACAGAGCACAATCACTCAAAATGAGCAAAACCAAGTCCCCCTCG	1140
Db	7020	ACCACCAATGACACACACAGAGCACAATCACTCAAAATGAGCAAAACCAAGTCCCCCTCG	7079
Qy	1141	ATAACAACTCTGCTCCAAAGGGAAATCCCAACACAGCAAAAGAGCACAGCAGCAAAAA	1200
Db	7080	ATAACAACTCTGCTCCAAAGGGAAATCCCAACACAGCAAAAGAGCACAGCAGCAAAAA	7139
Qy	1201	GGCCCCGCCACAAAGCCACCAACACAGCAAAATGAGCAATTCACAGTCTCCCTCCCAAC	1260
Db	7140	GGCCCCGCCACAAAGCCACCAACACAGCAAAATGAGCAATTCACAGTCTCCCTCCCAAC	7199
Qy	1261	CCAGCTCGACTGACACAACTCTTGTATATTCAGAGAAAGCGATCGATTTCTGGAAA	1320
Db	7200	CCAGCTCGACTGACACAACTCTTGTATATTCAGAGAAAGCGAGTATCTCTGGAGG	7259
Qy	1321	GAAGGTGATATATTCCTGTTTATAGATGGGTAAATAATCTGAAATTTGATTTGATCCA	1380
Db	7260	GAAGGTGATATATTCCTGTTTATAGATGGGTAAATAATCTGAAATTTGATTTGATCCA	7319
Qy	1381	ATCCAAACACAGAAACAACTTTTGTATGATGAAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
Db	7320	ATCCAAACACAGAAACAACTTTTGTATGATGAAATCTCCAGCTTTAATACTTCAACTAATGAG	7379
Qy	1441	GAACACACACTCCCGGAAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA	1500
Db	7380	GATCAACATGCTCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATAATAATGAG	7439
Qy	1501	GATATGCTACTCTCGGGAAACAGAGATGATTTGATGACAGATTTGAGGATTTGGAGT	1560
Db	7440	AACACTGCTACTCTCGGGAAATGAGAAATGATTTGATGACAGATTTAAGAAATTTGGAGC	7499
Qy	1561	GTGAGAGGAGAGATTTGGCGGACAGGCTTAGCTGGATACATTTTGGCCCTGGGAATC	1620
Db	7500	GTTCAGGAGGATGACTGCGCGAGGCTCAGTTGGATACCGTTTGGCCCTGGGAAT	7559
Qy	1621	GAAGGACTTATATCTGCGGTTTAAATCAAAATCAGAACTTATTTAGTTTGGTTGAGG	1680
Db	7560	GAAGGACTTATATCTGCGTGTTTTAAATCAAAATCAAAATTTGGTCTGCGGTTGAGG	7619
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTTGTAAGGGTCAACACCGAGGAA	1740
Db	7620	CGCTTAGCTAATCAAACTGCGCAATCTTGGAACTCTTATTTGAGAGTCACTGAGGAA	7679

Qy	1741	AGACATTTTCCTTAATCAATAGCATGCAATTTGACTTTTGCCTACGAGGTGGGGCGGA	1800
Db	7680	AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAACAATGGGAGGA	7739
Qy	1801	ACATGCAAGTGTAGGACCTGATTGTCATAGGAATAGAAGTCTATCTAAATAATC	1860
Db	7740	ACATGCAAAGTGTGTGACCTGATTGTTGCATCGGATAGAAGACTTGTCCAAAAATATT	7799
Qy	1861	TCAGAAACAAATGCACAAAAATCAGAAAGGATGAACAAAAGGAGGAAACTGCTCGGGGTCTA	1920
Db	7800	TCAGAGCAAAATGACCAAAATTAANAAGGACGAACAAAAGAGGGGACTTGTGTGGGGTCTG	7859
Qy	1921	GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTTCACCAATTTGGGCATCTCGTACTA	1980
Db	7860	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGCTACTA	7919
Qy	1981	TTATCTATAGCTGTCTGATTGCTGTCTCTGTATCTGTCTGATCTTTCACTAAATACATT	2040
Db	7920	TTATCCATAGCTGTCTTGATTGCTCTATCCTGTATTTGTCGTATCTTTACTAAATATATC	7979
Qy	2041	GGATGA	2046
Db	7980	GGATAA	7985

RESULT 3
US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
;

[illegible]

Db 1057 GAGCAATGGNACCGACCCACACACACACACAGCAAGCTCCTCAACAATACTAATACA 1116
Qy 1021 ATCTCTACTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAC 1080
Db 1117 ACTCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCAAACCGCAAC 1176
Qy 1081 ACCACCATGACACACAGCAGCAGCATCTGAAATGAGCAACAGTGCCCTCCGCTCG 1140
Db 1177 ATCCACCAATATGATACACACGTCGCAACTAGCAGAAAGCAGCAACCAATGCTCAGTTG 1236
Qy 1141 ATAACAACCCCTCCCTCCAAACGGGAAATCCCAACACAGCAAGAGACACAGCAACAAAA 1200
Db 1237 AACACAATCTAGATCCAAAGAGAAATCCCAACACAGCAGCAGCAACACAGCAAC 1296
Qy 1201 GGGCCCGCACAAAGGCAACCAACACGACAAATGAGCATTTTCAACGCTCTCCGCCAC 1260
Db 1297 AACATCATCATGACATCATGATATACAAAGCAACACCCCAACAATTTCTCTCGGAT 1356
Qy 1261 CCCAGCTGACGACACATCTTGTATATTTTCAGAGAGAGCGATCGATTTCTGGAA 1320
Db 1357 TCTAGTCGCAACCCGCCCTCTATATCTTTAGAAAGAACGAGCAATTTCTCTGGAA 1416
Qy 1321 GAAGTGATATATTCCTCCCTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Db 1417 GAAGTGATATATTCCTCCCTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1476
Qy 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db 1477 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1536
Qy 1441 GAACACACACTCCCCCGAATATCAGTTTAATCTTCTTCTTATTTCTTGATAAAAATGGA 1500
Db 1537 GAACACACACTCCCCCGAATATCAGTTTAATCTTCTTCTTATTTCTTGATAAAAATGGA 1596
Qy 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGATGACAGAGTTTTCAGGATTTGGAGT 1560
Db 1597 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGATGACAGAGTTTTCAGGATTTGGAGT 1656
Qy 1561 GTGAGGAGGACGATTTTGGCGGAGGGCTTACGTGATACCAATTTTGGCCCTGGAATC 1620
Db 1657 GTGAGGAGGACGATTTTGGCGGAGGGCTTACGTGATACCAATTTTGGCCCTGGAATC 1716
Qy 1621 GAAGGACTCTACTGCGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Db 1717 GAAGGACTCTACTGCGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1776
Qy 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAGGCTCACAACCGAGAA 1740
Db 1777 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAGGCTCACAACCGAGAA 1836
Qy 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGATCTTTTGTCTTACGAGGTGGGCGGA 1800
Db 1837 AGGACATTTTCTTAATCAATAGGATGCAATTTGATCTTTTGTCTTACGAGGTGGGCGGA 1896
Qy 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 1897 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1956
Qy 1861 TCAGAAACAAATCGAACAATCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGCTCTA 1920
Db 1957 TCAGAAACAAATCGAACAATCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGCTCTA 2016
Qy 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCATCTCTGCTACTA 1980
Db 2017 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCATCTCTGCTACTA 2076
Qy 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGATCTCTGCTGATCTTCACTAAATACAT 2040
Db 2077 TTATCTATAGCTGTTCTGATTTGCTCTGCTGATCTCTGCTGATCTTCACTAAATACAT 2136
Qy 2041 GGATGA 2046
Db 2137 GGATGA 2142

RESULT 4
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172
; OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1
Query Match 6.4%; Score 131.2; DB 3; Length 2172;
Best Local Similarity 61.9%; Pred. No. 7.4e-30;
Matches 208; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
Qy 1555 TGGAGTGTGCGAGGAGGAGGATTTGCGCGCAGGCGCTAGCTGATACCAATTTTGGCCCT 1614
Db 1693 TGGAGTGTGCGAGGAGGAGGATTTGCGCGCAGGCGCTAGCTGATACCAATTTTGGCCCA 1752
Qy 1615 GGAATCGAAGGACTCTATCTATCTGCGGCTTTTAAATCAAAATCAGAAACAATTTAGTTGTAGG 1674
Db 1753 GCAGCCGAGGAAATTTACATAGAGGGGCTAATGCACAAATCAAGATGGTTAATCTGTGG 1812
Qy 1675 TTGAGGCGCTTAGCTAATCAAAATCGCTTAAATCTTGGAGCTCTTGTTAAGGGTCAACAAC 1734
Db 1813 TTGAGACAGCTGGCGCAACGAGACGACTCAAGCTCTTCAACTGTCTCTGAGAGCCACAAC 1872
Qy 1735 GAGGAAGGACATTTTCTTTAATCAATAGGATCAATTTGACTTTTCTTACGAGGTGG 1794
Db 1873 GAGCTAGCACCTTTTCAATCTCAACCGTAAGCAATTTGATTTCTTCTGCTGACGATGG 1932

```
QY 1795 GCGGAAACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAAA 1854
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1933 GCGGCACATGCCACATCTTGGACCGGACTGCTGTATCGAACCACATGATGGACCAAG 1992
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1855 AATATCTCAGACRAATCGACAAATCAGAAAGGAT 1890
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1993 AACATAACAGACAAAATTGATCAGATTATTCATGAT 2028
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-650-086A-1
; Sequence 1, Application US/09650086A
; Patent No. 6630144
; GENERAL INFORMATION:
; APPLICANT: Hart, Mary Kate
; APPLICANT: Wilson, Julie A.
; APPLICANT: Schmalohn, Alan L.
; TITLE OF INVENTION: Prophylactic and Therapeutic Monoclonal Antibodies
; FILE REFERENCE: 003/198/SAP
; CURRENT APPLICATION NUMBER: US/09/650,086A
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: US 60/151,505
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola Zaire
US-09-650-086A-1

Query Match 6.4%; Score 131.2; DB 4; Length 2298;
Best Local Similarity 61.9%; Pred. No. 7.7e-30;
Matches 208; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1555 TGAAGTGTGCAGAGGACGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCT 1614
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1610 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTGGATACCATATTTTCGGGCA 1669
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1615 GGAATCGAGGACTCTATCTGCGGTTTATCAAAATCAGAAATTTAGTTTGTAGG 1674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1670 GCAGCCGAGGGAATTTACATAGAGGGGCTAATGCAATCAAGATGGTTTATCTGTGGG 1729
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1675 TTGAGCGGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCACAACC 1734
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1730 TTGAGACAGCTGGCCCAACGAGAGGACTCAGCTCTTCACTGTTCTTGAGGCCAAT 1789
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1735 GAGGAAGACATTTTCTTAATCAATAGGCAATGCAATGACTTTTGTACGAGGTGG 1794
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1790 GAGTACGACACCTTTTCAATCTCAACCGTAAGGCAATGATTTCTGTGCGCGATGG 1849
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1795 GCGGNAACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAAA 1854
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1850 GCGGCACATGCCACATCTTGGGACCGGACTCTGTATCGAACCACATGATGGACCAAG 1909
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1855 AATATCTCAGAACAAATCAGAAATCAGAAAGGAT 1890
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1910 AACATAACAGACAAAATTGATCAGATTATTCATGAT 1945
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-949-016-13845/c
; Sequence 13845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13845
; LENGTH: 152132
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13845

Query Match 3.6%; Score 74.6; DB 4; Length 152132;
Best Local Similarity 50.6%; Pred. No. 8.1e-11;
Matches 206; Conservative 0; Mismatches 199; Indels 2; Gaps 1;

QY 872 CATCAACAAATGCCACCCACTCCCTCACCACACCAAGCAGCCGACGAGGAGAAACA 931
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136819 CATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136760
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 932 ACACAAACCAATTCCTGACTGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 991
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136759 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136700
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 992 CGT--CGATGCCCTCATACACTACACAAATCTTACTTAACAACTCTCCAAACACAA 1049
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136699 CATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136640
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1050 CTTGAGCACTCTCTGCAACCAATTAACAAACCAACCAACCAACCAACCAACCAAC 1109
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136639 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136580
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1110 CACTGAAATGAGCAAAACAGTGCCTCCCTCGATTAACAACTCTCCAAAGGGAAATCC 1169
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136579 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1170 CACCACGAAAGAGCCAGCAGCAAAAAGGCGCCGCAACGCGCAACCAACCAACCAAC 1229
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136519 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136460
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1230 AATGAGCAATTTACCAGTCTCTCCCAACCAACCAACCAACCAACCAACCAACCAAC 1276
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136459 CACCATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136413
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-949-016-12371/c
; Sequence 12371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12371
; LENGTH: 152145
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12371

Query Match 3.6%; Score 74.6; DB 4; Length 152145;
Best Local Similarity 50.6%; Pred. No. 8.1e-11;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1285.84 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacatgttctct.....tcactaaatcatgtgatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	15	US-10-066-506A-5
2	1841.2	90.0	7778	20	US-10-491-121-30
3	1841.2	90.0	11460	16	US-10-267-322-1
4	1813.6	88.6	2051	15	US-10-066-506A-9
5	1767.6	86.4	7005	20	US-10-491-121-31
6	1767.6	86.4	8256	20	US-10-491-121-32
7	1692.4	82.7	2046	21	US-10-811-353-2
8	1692.4	82.7	2046	15	US-10-066-506A-5
9	1518	74.2	2046	15	US-10-066-506A-13
10	1319.6	64.5	2046	15	US-10-066-506A-7
11	1308	63.9	1841	15	US-10-066-506A-3
12	995.6	48.7	6902	20	US-10-491-121-42
13	546.6	26.7	2252	15	US-10-066-506A-1
14	145	7.1	6940	20	US-10-491-121-15
15	145	7.1	7073	22	US-10-860-878-2
16	145	7.1	7082	20	US-10-491-121-13
17	145	7.1	7087	20	US-10-491-121-14
18	145	7.1	8221	20	US-10-491-121-27
19	145	7.1	8338	20	US-10-491-121-26
20	145	7.1	10783	20	US-10-491-121-25
21	138.8	6.8	6885	20	US-10-491-121-18
22	138.8	6.8	6889	20	US-10-491-121-19
23	138.8	6.8	7002	20	US-10-491-121-16
24	138.8	6.8	7003	22	US-10-860-878-1
25	138.8	6.8	7023	20	US-10-491-121-21
26	138.8	6.8	7036	20	US-10-491-121-17
27	138.8	6.8	8146	20	US-10-491-121-20
28	131.2	6.4	2039	15	US-10-066-506A-11
29	131.2	6.4	2298	9	US-09-337-946A-1
30	131.2	6.4	2298	17	US-10-384-976-1
31	131.2	6.4	2298	18	US-10-226-795-1
32	131.2	6.4	2298	22	US-10-696-633-1
33	131.2	6.4	4667	20	US-10-491-121-10
34	131.2	6.4	6561	20	US-10-491-121-4
35	131.2	6.4	6624	20	US-10-491-121-3
36	131.2	6.4	6914	20	US-10-491-121-9
37	131.2	6.4	7154	20	US-10-491-121-1
38	131.2	6.4	7188	20	US-10-491-121-2
39	131.2	6.4	7272	22	US-10-860-878-4
40	131.2	6.4	7285	22	US-10-860-878-3
41	131.2	6.4	8199	20	US-10-491-121-29
42	131.2	6.4	8439	20	US-10-491-121-28
43	131.2	6.4	18959	17	US-10-353-856-19
44	131.2	6.4	18959	17	US-10-353-856-37
45	131.2	6.4	18959	17	US-10-353-856-46

8	1692.4	82.7	19112	17	US-10-353-856-27	Sequence 27, Appl
9	1518	74.2	2046	15	US-10-066-506A-13	Sequence 13, Appl
10	1319.6	64.5	2046	15	US-10-066-506A-7	Sequence 7, Appl
11	1308	63.9	1841	15	US-10-066-506A-3	Sequence 3, Appl
12	995.6	48.7	6902	20	US-10-491-121-42	Sequence 42, Appl
13	546.6	26.7	2252	15	US-10-066-506A-1	Sequence 1, Appl
14	145	7.1	6940	20	US-10-491-121-15	Sequence 15, Appl
15	145	7.1	7073	22	US-10-860-878-2	Sequence 2, Appl
16	145	7.1	7082	20	US-10-491-121-13	Sequence 13, Appl
17	145	7.1	7087	20	US-10-491-121-14	Sequence 14, Appl
18	145	7.1	8221	20	US-10-491-121-27	Sequence 27, Appl
19	145	7.1	8338	20	US-10-491-121-26	Sequence 26, Appl
20	145	7.1	10783	20	US-10-491-121-25	Sequence 25, Appl
21	138.8	6.8	6885	20	US-10-491-121-18	Sequence 18, Appl
22	138.8	6.8	6889	20	US-10-491-121-19	Sequence 19, Appl
23	138.8	6.8	7002	20	US-10-491-121-16	Sequence 16, Appl
24	138.8	6.8	7003	22	US-10-860-878-1	Sequence 1, Appl
25	138.8	6.8	7023	20	US-10-491-121-21	Sequence 21, Appl
26	138.8	6.8	7036	20	US-10-491-121-17	Sequence 17, Appl
27	138.8	6.8	8146	20	US-10-491-121-20	Sequence 20, Appl
28	131.2	6.4	2039	15	US-10-066-506A-11	Sequence 11, Appl
29	131.2	6.4	2298	9	US-09-337-946A-1	Sequence 1, Appl
30	131.2	6.4	2298	17	US-10-384-976-1	Sequence 1, Appl
31	131.2	6.4	2298	18	US-10-226-795-1	Sequence 1, Appl
32	131.2	6.4	2298	22	US-10-696-633-1	Sequence 1, Appl
33	131.2	6.4	4667	20	US-10-491-121-10	Sequence 10, Appl
34	131.2	6.4	6561	20	US-10-491-121-4	Sequence 4, Appl
35	131.2	6.4	6624	20	US-10-491-121-3	Sequence 3, Appl
36	131.2	6.4	6914	20	US-10-491-121-9	Sequence 9, Appl
37	131.2	6.4	7154	20	US-10-491-121-1	Sequence 1, Appl
38	131.2	6.4	7188	20	US-10-491-121-2	Sequence 2, Appl
39	131.2	6.4	7272	22	US-10-860-878-4	Sequence 4, Appl
40	131.2	6.4	7285	22	US-10-860-878-3	Sequence 3, Appl
41	131.2	6.4	8199	20	US-10-491-121-29	Sequence 29, Appl
42	131.2	6.4	8439	20	US-10-491-121-28	Sequence 28, Appl
43	131.2	6.4	18959	17	US-10-353-856-19	Sequence 19, Appl
44	131.2	6.4	18959	17	US-10-353-856-37	Sequence 37, Appl
45	131.2	6.4	18959	17	US-10-353-856-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-10-066-506A-5
; Sequence 5, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh_Microsoft Word 6.0
; SEQ ID NO 5
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 2
; OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 100.0%; Score 2046; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTTCCTTATCATGCTTATCTATTAATTCAGGGACAAAATCTCC 60

Db 1 ATGAAGACCATGTTTCCTTATAGTCTTATCTTAAATTCAGGGACAAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
Db 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
QY 121 CTCGAGAGACAGAAAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGAGACAGAAAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
Db 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
QY 241 GTTGTAGTACAGAGGGGGAGGAGCCCAAAACATGCTCAATATAGTGTAAACGATCCC 300
Db 241 GTTGTAGTACAGAGGGGGAGGAGCCCAAAACATGCTCAATATAGTGTAAACGATCCC 300
QY 301 TCTGGAATAATCCTTGCTGTAGATCCTCTACCAACATCCCGTGAATATCCGAAATGCAAA 360
Db 301 TCTGGAATAATCCTTGCTGTAGATCCTCTACCAACATCCCGTGAATATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTCAAAGGTCAAAACCTCTATGACACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTCAAAGGTCAAAACCTCTATGACACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACTGTACCGAGGGCAAGTCTTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCACAACTGTACCGAGGGCAAGTCTTC 480
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAATGATTTTCTCG 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAATGATTTTCTCG 540
QY 541 CGGCAAGGCAAGGGTACCGTCATATGAATCTGAATCTTCTACTAATAAATATTTGGCAAGT 600
Db 541 CGGCAAGGCAAGGGTACCGTCATATGAATCTGAATCTTCTACTAATAAATATTTGGCAAGT 600
QY 601 AGTACGGNACCAACAGATGACACTGGATGTTTGGGGCTCTTCNAGNATACAAATCT 660
Db 601 AGTACGGNACCAACAGATGACACTGGATGTTTGGGGCTCTTCNAGNATACAAATCT 660
QY 661 ACAAGAACCAACATGTCTCCGTCCAAATACCTCCACACTGCACACGCCCGTCCG 720
Db 661 ACAAGAACCAACATGTCTCCGTCCAAATACCTCCACACTGCACACGCCCGTCCG 720
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCAAGC 780
Db 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCAAGC 780
QY 781 AGTGATGATGAGGACCTCGCAATCCGGCTCAGGGTCGGAGACGAGAACCCACACA 840
Db 781 AGTGATGATGAGGACCTCGCAATCCGGCTCAGGGTCGGAGACGAGAACCCACACA 840
QY 841 ACTTCTGATCGGTCAACCAAGCAAGGGCTTTCATCAACAAATGCCACCACTCCCTCACA 900
Db 841 ACTTCTGATCGGTCAACCAAGCAAGGGCTTTCATCAACAAATGCCACCACTCCCTCACA 900
QY 901 CAACCAAGCACGCCACAGCAAGGAGAAACAAACAAACCATTTCCCAAGATCTGTGACT 960
Db 901 CAACCAAGCACGCCACAGCAAGGAGAAACAAACAAACCATTTCCCAAGATCTGTGACT 960
QY 961 GAACTAGACAAAATTAACACACTGCACACCGTCCATGCCCCCTCATTAACACTACCACA 1020
Db 961 GAACTAGACAAAATTAACACACTGCACACCGTCCATGCCCCCTCATTAACACTACCACA 1020
QY 1021 ATCTCTACTAACAACACCTCCAAACCAACTTTCAGCACTCTCTCTGCACCAATTAACAAAC 1080
Db 1021 ATCTCTACTAACAACACCTCCAAACCAACTTTCAGCACTCTCTCTGCACCAATTAACAAAC 1080
QY 1081 ACCACCAATGACACACAGAGCAAAATCACTGAAAAATGAGCAAAACCAAGTGCCCCCTCG 1140

Db 1081 ACCACCAATGACACACAGAGCAAAATCACTGAAAAATGAGCAAAACCAAGTGCCCCCTCG 1140
QY 1141 ATAACAACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGACCAACAGAGCAAAAAA 1200
Db 1141 ATAACAACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGACCAACAGAGCAAAAAA 1200
QY 1201 GGCCCCGGCAACAAGGGCACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCACCC 1260
Db 1201 GGCCCCGGCAACAAGGGCACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCACCC 1260
QY 1261 CCCAGCTCGATGTCACACATCTTGTATATTTTCAGAGAAAGAGCGATCGATTTCTGGAAA 1320
Db 1261 CCCAGCTCGATGTCACACATCTTGTATATTTTCAGAGAAAGAGCGATCGATTTCTGGAAA 1320
QY 1321 GAAAGTGATATATTTCCCGTTTTAGATGGGTTAAATAAATACTGAAATTTGATTTTTCATCCA 1380
Db 1321 GAAAGTGATATATTTCCCGTTTTAGATGGGTTAAATAAATACTGAAATTTGATTTTTCATCCA 1380
QY 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAAATACTTCAACTAATGAG 1440
Db 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAAATACTTCAACTAATGAG 1440
QY 1441 GAACACACACTCCCGCGAATATCAGTTTAACTTCTCTTATTTTCTCTGATAAAAAATGGA 1500
Db 1441 GAACACACACTCCCGCGAATATCAGTTTAACTTCTCTTATTTTCTCTGATAAAAAATGGA 1500
QY 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGTGATGAGAGTTGAGGATTTGGAGT 1560
Db 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGTGATGAGAGTTGAGGATTTGGAGT 1560
QY 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1620
Db 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1620
QY 1621 GAAGACTCTATCTGCTCGGTTTTAAATCAAAAATCAGAACTTTAGTTGTAGGTTGAGG 1680
Db 1621 GAAGACTCTATCTGCTCGGTTTTAAATCAAAAATCAGAACTTTAGTTGTAGGTTGAGG 1680
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCAACACCGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCAACACCGAGGAA 1740
QY 1741 AGGACATTTTCTTAAATCAATAGGATGCAATGATGCTTTTGTCTTACGAGGTGGGGCGGA 1800
Db 1741 AGGACATTTTCTTAAATCAATAGGATGCAATGATGCTTTTGTCTTACGAGGTGGGGCGGA 1800
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
QY 1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCCACCAATTTGGGCATCTCTGTACTA 1980
Db 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCCACCAATTTGGGCATCTCTGTACTA 1980
QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGTATCTGCTGATCTTCACTAAATACATT 2040
Db 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGTATCTGCTGATCTTCACTAAATACATT 2040
QY 2041 GGATGA 2046
Db 2041 GGATGA 2046

RESULT 2
US-10-491-121-30
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY

APPLICANT: YANG, ZHI-YONG
APPLICANT: SULLIVAN, NANCY
APPLICANT: SANCHEZ, ANTHONY
TITLE OF INVENTION: Development of a Preventive Vaccine for
FILE OF INVENTION: Filovirus Infection in Primates
FILE REFERENCE: NIH221.001NP
CURRENT APPLICATION NUMBER: US/10/491,121
CURRENT FILING DATE: 2004-03-26
PRIOR APPLICATION NUMBER: PCT/US02/30251
PRIOR FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: US 60/326476
PRIOR FILING DATE: 2001-10-01
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 7778
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30

Query Match 90.0%; Score 1841.2; DB 20; Length 7778;

Best Local Similarity 93.7%; Pred. No. 0;

Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY	1	ATGAAGACACATGTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC	60
DB	2034	ATGAAGACACATGTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC	2093
QY	61	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGTCGCGAACT	120
DB	2094	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCGGTATGTCGCGAACT	2153
QY	121	CTCCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180
DB	2154	CTCCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	2213
QY	181	GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
DB	2214	GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	2273
QY	241	GTTGAGTACACAGGGGGGAGGAGCCAAACATGCTACAATATAGGTAAACGGATCCC	300
DB	2274	GTTGAGTACACAGGGGGGAGGAGCCAAACATGCTACAATATAGGTAAACGGATCCC	2333
QY	301	TCTGAAAAATCTTCTGTTAGATCCTCTACCAACATCCGAGACTATCCGAAATGCAAA	360
DB	2334	TCTGAAAAATCTTCTGTTAGATCCTCTACCAACATCCGAGACTATCCGAAATGCAAA	2393
QY	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
DB	2394	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	2453
QY	421	GGAGCATTTTCTGTATGATCGATTCCTCCACAAATGTACCGAGGCAAGTCTTC	480
DB	2454	GGAGCATTTTCTGTATGATCGATTCCTCCACAAATGTACCGAGGCAAGTCTTC	2513
QY	481	ACTGAAGGAAACATAGCAGTATGATGTCATAAGACAGTGCACAAAATGATTTTCTCG	540
DB	2514	ACTGAAGGAAACATAGCAGTATGATGTCATAAGACAGTGCACAAAATGATTTTCTCG	2573
QY	541	CGGCAAGAACAGGGTACCGTCAATATGATCTGACTTCTTACTAATAAATATTTGGACAAGT	600
DB	2574	CGGCAAGAACAGGGTACCGTCAATATGATCTGACTTCTTACTAATAAATATTTGGACAAGT	2633
QY	601	AGTAACGGAAACGAAACGAAATGACACTGATGTTTGGGGCTCTTCAAGAAATACAAATCT	660
DB	2634	AGTAACGGAAACGAAACGAAATGACACTGATGTTTGGGGCTCTTCAAGAAATACAAATCT	2693
QY	661	ACAAAGAACCAACATGTCTCGGTCCAAAATACCTCCACCACTGCCACACCGCGTCCG	720
DB	2694	ACAAAGAACCAACATGTCTCGGTCCAAAATACCTCCACCACTGCCACACCGCGTCCG	2753

QY	721	GAGATCAAACTCAACAGACCCCAACTGATGCGCACCAAACTCAATACACGACCCCAAGC	780
DB	2754	GAGATCAAACTCAACAGACCCCAACTGATGCGCACCAAACTCAATACACGACCCCAAGC	2813
QY	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	840
DB	2814	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	2873
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCGCACCCCTCCCTCACA	900
DB	2874	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCGCACCCCTCCCTCACA	2933
QY	901	CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAAGATGCTGTGACT	960
DB	2934	CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAAGATGCTGTGACT	2993
QY	961	GAACTAGACAAAAATAACACAACTGCAACACCGTCCATGCCCCCTCATATAACTACCA	1020
DB	2994	GAACTAGACAAAAATAACACAACTGCAACACCGTCCATGCCCCCTCATATAACTACCA	3053
QY	1021	ATCTCTACTAAACAACCTCCANAACAACATTTAGCACTCTCTGTGACCACTTACAAAC	1080
DB	3054	ATCTCTACTAAACAACCTCCANAACAACATTTAGCACTCTCTGTGACCACTTACAAAC	3113
QY	1081	ACCAACATGACACACACAGACACATCACTGAAATGAGCAACACAGTGCCTCCCTCG	1140
DB	3114	ACCAACATGACACACACAGACACATCACTGAAATGAGCAACACAGTGCCTCCCTCG	3173
QY	1141	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA	1200
DB	3174	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA	3233
QY	1201	GGCCCCGACACAGCGCACCAACACGACAAATGAGCAATTTTCCAGCTCTCCCTCCACC	1260
DB	3234	GGCCCCGACACAGCGCACCAACACGACAAATGAGCAATTTTCCAGCTCTCCCTCCACC	3293
QY	1261	CCAGCTCGACTGCAACAACATCTTGTATATTTTCAAGAAAGCGATCGATTTCTGGAA	1320
DB	3294	CCAGCTCGACTGCAACAACATCTTGTATATTTTCAAGAAAGCGATCGATTTCTGGAG	3353
QY	1321	GAACTGATATATTTCCCGTTTTAGATGGGTTAATAAATACTGAAATTTGATTTTCATCA	1380
DB	3354	GAACTGATATATTTCCCGTTTTAGATGGGTTAATAAATACTGAAATTTGATTTTCATCA	3413
QY	1381	ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
DB	3414	ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG	3473
QY	1441	GAAACAACAATCCCGCGCAATATCAGTTTAACTTTCTCTTATTTTCTCTGATAAAAATGA	1500
DB	3474	GATCAACAATGCTCCCGCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATAATAAG	3533
QY	1501	GATCTGCTACTCTGGGGAACGAGAAATGATTTGATGATGAGTTCAGAGTTTGGAGT	1560
DB	3534	AACATGCTACTCTGGGGAACGAGAAATGAGAAATGATTTGATGATGAGTTCAGAGTTTGGAG	3593
QY	1561	GTGAGGAGGACGATTTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC	1620
DB	3594	GTTGAGGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTTTGGCCCTCGAATC	3653
QY	1621	GAAGGACTCTATCTGCGGTTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGTTGAGG	1680
DB	3654	GAAGGACTCTATCTGCTGTTTTAATTAATAAATAAATAAATTTGCTGCAAGTTGAGG	3713
QY	1681	CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCACACCCAGGA	1740
DB	3714	CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCACACCCAGGA	3773
QY	1741	AGGACATTTCTCTTAATCAATAGGATCAATTTGATTTTGTCTTACGAGGTGGCGGGA	1800
DB	3774	AGGACATTTCTCTTAATCAATAGGATCAATTTGATTTTGTCTTACGAGGTGGCGGGA	3833

```
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAATAATATC 1860
Db 3834 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTGTCACAAATAATTT 3893
QY 1861 TCAGAACAAATCGACAAATCAGAAAGTGAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAANAAGGACGAAACAAAGAGGGGACTGGTTGGGGTCTG 3953
QY 1921 GGTGGCAAAATGGTGACATCTGACTGGGGTGTCTCACCAAATTTGGGCATCTCTGCTACTA 1980
Db 3954 GGTGTAATGTTGGTACATCCGACTGGGGTGTCTTACTACTGCGGATTTTGGTACTA 4013
QY 1981 TTAATCTATAGCTGTTCTGATGCTCTGCTCTGCTATCTGTCGATCTTCACTAAATACATTT 2040
Db 4014 TTATCCATAGCTGTTCTGATGCTCTATCTCTGATTTGTCGATCTTTACTAAATATATC 4073
QY 2041 GGATGA 2046
Db 4074 GGATAA 4079

RESULT 3
US-10-267-322-1
; Sequence 1, Application US/10267322
; Publication No. US20030152590A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143D
; CURRENT APPLICATION NUMBER: US/10/267,322
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/336,910, and US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-10-267-322-1

Query Match 90.0%; Score 1841.2; DB 16; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAATAATCTCCCC 60
Db 5940 ATGAGACACATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAATAATCTCCCC 5999
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATCGGTATGCTCCCGAACT 120
Db 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATCGGTATGCTCCCGAACT 6059
QY 121 CTCGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAT 240
Db 6120 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAT 6179
QY 241 GTTGTAGTACAGAGGGGGAGGAGCAAAAATGCTACAAATATAGTGAACGATCC 300
Db 6180 GTTGTAGTACAGAGGGGGAGGAGCAAAAATGCTACAAATATAGTGAACGATCC 6239
QY 301 TCTGGAAAATCTTGCTGTAGATCTCTTACCAATCCGTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAAAATCTTGCTGTAGATCTCTTACCAATCCGTGACTATCCGAAATGCAAA 360
```

```
Db 6240 TCTGGAAAATCTTGCTGTAGATCTCTTACCAATCCGTGACTATCTTAAATGCAAA 6299
QY 361 ACTATCCATCATATTTCAAGGTCAAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGGATTCGCTCCACAAATGTATCCGAGGCAAAATGCTCTTC 480
Db 6360 GGAGCATTTTTTCTGTATGATCGGATTCGCTCCACAAATGTATCCGAGGCAAAATGCTCTTC 6419
QY 481 ACTGAAAGGACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAATGATTTTCTTCG 540
Db 6420 ACTGAAAGGACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAATGATTTTCTTCG 6479
QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTTACTATAAATATTTGGACAAGT 600
Db 6480 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTTACTATAAATATTTGGACAAGT 6539
QY 601 AGTAAACGGAACGCAAAACGAATGACATGGATGTTTGGCGCTCTTTCAAGAATATCAATTTCT 660
Db 6540 AGTAAACGGAACGCAAAACGAATGACATGGATGTTTGGCGCTCTTTCAAGAATATCAATTTCT 6599
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCACTGCCACACAGCCCGTCCG 720
Db 6600 ACAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCACTGCCACACAGCCCGTCCG 6659
QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 780
Db 6660 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 6719
QY 781 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACACA 840
Db 6720 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACACA 6779
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db 6780 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 6839
QY 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 6899
QY 961 GAACCTAGACAAATAATACAACTGACAAACCGTCCATGCCCCCTCATAACTACCACA 1020
Db 6900 GAACCTAGACAAATAATACAACTGACAAACCGTCCATGCCCCCTCATAACTACCACA 6959
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACAAAC 1080
Db 6960 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACAAAC 7019
QY 1081 ACCACCAATGACACACAGAGCAGCAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 1140
Db 7020 ACCACCAATGACACACAGAGCAGCAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 7079
QY 1141 ATAAACACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGAGCAAAAAA 1200
Db 7080 ATAAACACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGAGCAAAAAA 7139
QY 1201 GGGCCCGCACACAGGACCAAAACAGCAAAATGAGCATTTTCCAGCTCTCCCGCCAC 1260
Db 7140 GGGCCCGCACACAGGACCAAAACAGCAAAATGAGCATTTTCCAGCTCTCCCGCCAC 7199
QY 1261 CCCAGCTCGACTGCAACAACATCTTGTATATTTTCAAGAAAGCGATCGATTTTCTGGAAA 1320
Db 7200 CCCAGCTCGACTGCAACAACATCTTGTATATTTTCAAGAAAGCGAGTATCTCTTGGAGG 7259
QY 1321 GAAGGTGATATATTTCCCGTTTTTGTAGATGGTTTAAATACTGAAATGATTTTATCA 1380
Db 7260 GAAGGTGATATATTTCCCGTTTTTGTAGATGGTTTAAATACTGAAATGATTTTATCA 7319
QY 1381 ATCCCAACACAGAAACATCTTTGATGATCTCCAGCTTTTAACTCACTCAATATGAG 1440
Db 7320 GTTCCAAATACAAACCAATCTTTGATGATCTCTTAGTCTTGGTGGCTCGGCTGAGGAA 7379
```

```
QY 1441 GAAACAACACATCCCGGAAATATCAGTTTAACTTCTCTTATTTCTGATATAAAATGGA 1500
DB 7380 GATCAACATGCTCCCGCAATATAGTTTAACTTATCTTATTTCTTAAATATAATAGAG 7439
QY 1501 GATCTGCTCTCTCTGGGGAACACAGAAATGATTTGATGATGAGAGTTGAGAGT 1560
DB 7440 AACCTGCTCTCTCTGGGGAACATGAGAAATGATTTGATGATGAGAGTTGAGAGT 7499
QY 1561 GTGAGGAGGAGGATTTGGCGGAGGCTTACGCTGATACCAATTTTTGGCCCTGGATC 1620
DB 7500 GTTCAGGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGCCCTGGAT 7559
QY 1621 GAAGGACTTATATCTGGCGGTTTAAATCAAAATAGAAATATGATTTGATGAGT 1680
DB 7560 GAAGGACTTATACCTGCTGTTTAAATCAAAATAGAAATATGATTTGATGAGT 7619
QY 1681 CGCTTAGCTTAAATCAAAATGCTTAAATCTTTGGAGCTTTGTTAAAGGCTCACAACCGAGAA 1740
DB 7620 CTTCTAGCAATCAAAATGCTTAAATCTTTGGAGCTTTGTTAAAGGCTCACAACCGAGAA 7679
QY 1741 AGGACATTTCTTAAATCAATAGGATGCAATTTGATTTTGGCTTACGAGGTGGGCGGA 1800
DB 7680 AGAATCTTCTTAAATCAATAGATGCTTATTTGATTTTCTACTCACAAGATGGGAGGA 7739
QY 1801 ACATCAAGGCTCTAGGACCTGATTTGTTGATAGAAATAGAGATCTATCTAAATATC 1860
DB 7740 ACATCAAGGCTCTAGGACCTGATTTGTTGATAGAAATAGAGATCTATCTAAATATC 7799
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAAGGAGGAACTGGCTGGGCTCTA 1920
DB 7800 TCAGAACAAATCGACAAATCAGAAAGATGAAAGGAGGAACTGGCTGGGCTCTG 7859
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGCTGTTCTCACCAGTTGGGCTCTGCTACTA 1980
DB 7860 GGTGCAAAATGGTGACATCTGACTGGGCTGTTCTCACCAGTTGGGCTCTGCTACTA 7919
QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
DB 7920 TTATCTATAGCTGTTCTGATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7979
QY 2041 GGATGA 2046
DB 7980 GGATGA 7985
```

RESULT 4

```
; Sequence 9, Application US/10066506A
; Publication NO. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan. L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 9
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Marburg virus strain Musoke
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-9
```

Query Match 88.6%; Score 1813.6; DB 15; Length 2051;
Best Local Similarity 93.4%; Pred. No. 0;

```
Matches 1910; Conservative 0; Mismatches 124; Indels 12; Gaps 1;
QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAAATTCAGGGAACAAAAATCTCCCC 60
DB 10 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAAATTCAGGGAACAAAAATCTCCCC 69
QY 61 ATTTTGAAGATAGTCTAGTAAATCAACCCCAAAATGTTGATTCGGTATGCTCCGGAAT 120
DB 70 ATTTTGAAGATAGTCTAGTAAATCAACCCCAAAATGTTGATTCGGTATGCTCCGGAAT 129
QY 121 CTCAGGAAGACAGAAAGACGTCATCTGATGGGATTCACACTGAGTGGGCAAAAAATTTGCT 180
DB 130 CTCAGGAAGACAGAAAGACGTCATCTGATGGGATTCACACTGAGTGGGCAAAAAATTTGCT 189
QY 181 GATTCCCTTTGGAGGATCCAAAGCATCGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
DB 190 GATTCCCTTTGGAGGATCCAAAGCATCGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 249
QY 241 GTTCAGTACACAGAGGGGGAGGAGCCAAACATGCTACAAATATAAGTGTAAACGATCCC 300
DB 250 GTTCAGTACACAGAGGGGGAGGAGCCAAACATGCTACAAATATAAGTGTAAACGATCCC 309
QY 301 TCTGAAAAATCCTTGTGTTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
DB 310 TCTGAAAAATCCTTGTGTTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 369
QY 361 ACTATCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB 370 ACTATCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 429
QY 421 GGACATTTTTCTGTATGATCGCATTCGCTCCCAACAAATGTACCGAGGCAAAATCTTC 480
DB 430 GGACATTTTTCTGTATGATCGCATTCGCTCCCAACAAATGTACCGAGGCAAAATCTTC 477
QY 481 ACTGAAGGAACATAGCAGCTATGATTTGCTCAATAGACAGTGCACAAATGATTTCTCG 540
DB 478 ACTGAAGGAACATAGCAGCTATGATTTGCTCAATAGACAGTGCACAAATGATTTCTCG 537
QY 541 CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAATAATATTTGCAAGT 600
DB 538 CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAATAATATTTGCAAGT 597
QY 601 AGTAACGGAACGCAACGAATGACATGGAATGTTTCGGGCTCTTCAAGATATCAATTTCT 660
DB 598 AGTAACGGAACGCAACGAATGACATGGAATGTTTCGGGCTCTTCAAGATATCAATTTCT 657
QY 661 ACAAGGAACCAAAACATGCTCCGTCGCAAAATATCTCCACCACTGCCCCACAGCCGCTCG 720
DB 658 ACAAGGAACCAAAACATGCTCCGTCGCAAAATATCTCCACCACTGCCCCACAGCCGCTCG 717
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCAAGGACCAAGC 780
DB 718 GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCAAGGACCAAGC 777
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGCTCCGGAGGAGGAGGAGGAGGAGG 840
DB 778 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGCTCCGGAGGAGGAGGAGGAGGAGG 837
QY 841 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACA 900
DB 838 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACA 897
QY 901 CAACCAAGCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 898 CAACCAAGCAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
QY 961 GAATGAGCAAAAAATACACAACTGACAAACCGTCCATGCCCCCTCATACCACTACCA 1020
DB 958 GAATGAGCAAAAAATACACAACTGACAAACCGTCCATGCCCCCTCATACCACTACCA 1017
QY 1021 ATCTCTACTTAACACACCTCCAAACACAACTTTCAGCTCTCTCTGCACTTACCAAAAC 1080
DB 1018 ATCTCTACTTAACACACCTCCAAACACAACTTTCAGCTCTCTCTGCACTTACCAAAAC 1077
```


Qy	661	ACAAAGAAC	CAAA	CATGTGCTCCGTCCAA	AAATACCTCC	ACACATGCGCCACAGACCCCGTCCG	720
Db	2694	ACAAAGAAC	CAAAA	CATGTGCTCCGTCCAA	AAATACCTCCACCATGCGCCACAGACCCCGTCCG	2753	
Qy	721	GAGATCAAACT	CA	CAGCACC	CCCAACTGATGC	CAACCAAACTCAATACCGACGCCAAGC	780
Db	2754	GAGATCAAACT	CA	CAGCACCCCAACTGATGC	CAACCAAACTCAATACCGACGCCAAGC	2813	
Qy	781	AGTGATGATGAGGACCTCG	CAACATCTCGGCTCAGGGTCCG	GAGAAACAGAAACCCCA	CACA	840	
Db	2814	AGTGATGATGAGGACCTCG	CAACATCTCGGCTCAGGGTCCG	GAGAAACAGAAACCCCA	CACA	2873	
Qy	841	ACTTCTGATGCGGTTC	CAAGCAAGGGTTCATCA	ACAATGCCACCCATCTCCCTCA	ACCA	900	
Db	2874	ACTTCTGATGCGGTTC	CAAGCAAGGGTTCATCA	ACAATGCCACCCATCTCCCTCA	ACCA	2933	
Qy	901	CAACCAAGCAGCCAC	AGCAGAGGAGAA	CAACAACCAACCATTC	CCCAAGATGCTGTCACT	960	
Db	2934	CAACCAAGCAGCCAC	AGCAGAGGAGAA	CAACAACCAACCATTC	CCCAAGATGCTGTCACT	2993	
Qy	961	GAACTAGACAAAAATA	ACAACTGCGACAA	CCGTCCATGCGCCCTCAT	ATAACATACCACA	1020	
Db	2994	GAACTAGACAAAAATA	ACAACTGCGACAA	CCGTCCATGCGCCCTCAT	ATAACATACCACA	3053	
Qy	1021	ATCTCTATTAACA	CACCTCCAA	ACAACAATTCAGCACTCTCTCTG	CAACAATACCAAAAC	1080	
Db	3054	ATCTCTATTAACA	CACCTCCAA	ACAACAATTCAGCACTCTCTCTG	CAACAATACCAAAAC	3113	
Qy	1081	ACCACCAATGACAC	ACAGACACAACTCA	CTGAAAATGAGCAAA	CCAGTGC	1140	
Db	3114	ACCACCAATGACAC	ACAGACACAACTCA	CTGAAAATGAGCAAA	CCAGTGC	3173	
Qy	1141	ATAACAACCCCTGCC	CTCCAA	CGGGAAATCCCA	CCACACAGCAAAAGACAC	1200	
Db	3174	ATAACAACCCCTGCC	CTCCAA	CGGGAAATCCCA	CCACACAGCAAAAGACAC	3233	
Qy	1201	GGCCCCGCGCA	CAACGGCA	CCAAACA	CAAGCAATGAGCATTTCA	1260	
Db	3234	GGCCCCGCGCA	CAACGGCA	CCAAACA	CAAGCAATGAGCATTTCA	3293	
Qy	1261	CCAGCTCGACTG	CAACACATCTTGTA	TATTTACAGAAAGCGATCGATTTCTG	GA	1320	
Db	3294	CCAGCTCGACTG	CAACACATCTTGTA	TATTTACAGAAAGCGATCGATTTCTG	GA	3353	
Qy	1321	GAAGGTGATATATTTCC	GGTTTTAGATGGGTTA	TAAATAC	TGATAATCTGAAATTTGATTTGAT	1380	
Db	3354	GAAGGTGATATATTTCC	GGTTTTAGATGGGTTA	TAAATAC	TGATAATCTGAAATTTGATTTGAT	3413	
Qy	1381	ATCCAAACACAGAA	CAAACTTTGATGA	ATCTCCAGCTTTAATACTCA	ACTAATGAG	1440	
Db	3414	ATCCAAACACAGAA	CAAACTTTGATGA	ATCTCCAGCTTTAATACTCA	ACTAATGAG	3473	
Qy	1441	GAAACAACACATCTCC	CGGAATATCAGTTTAA	CTTTCTTATTTTCTGTATAAAATGGA		1500	
Db	3474	GATCAACAATGCTCTC	CGCCCAATATTAGTTTAA	CTTTATTTCTTAATATAAATGAG		3533	
Qy	1501	GATACTGCTACTCTG	GGGAAAA	CGAGAA	TGATTTGATGACAGAGTTTGA	1560	
Db	3534	AACACTGCTACTCTG	GAGAAATGAGAA	TGATTTGATGACAGAGTTTGA	1560	3593	
Qy	1561	GTGAGGAGGACGATTTG	CGCGCAGGGCTTAG	CTGGATACCAATTTTGGCCCTCGAATC	1620		
Db	3594	GTTCAGGAGGATGACCT	GGCGCAGGGCTCAGTT	TGGATACCGTTTTTGGCCCTCGAATC	1620	3653	
Qy	1621	GAAGGACTCTATATCTG	CGCGTTTAA	TCAAAATCAGAA	CAATTTAGTTTGTAGTTGAGG	1680	
Db	3654	GAAGGACTCTTAC	TGCTGTTTTAA	TAAAAATCAAA	CAATTTGGTCTG	3713	
Qy	1681	CGCTTAGCTAATCA	ACTGCTA	ATCTTGAGCTCTGTTTAA	GGGTCA	1740	
Db	3714	CGTCTAGCCAA	TCAACTG	CCAAATCTTG	NACTCTTATTTGAGATCA	3773	
Qy	1741	AGGACATTTTCT	TTAATCAATPAG	CACTGCAATTCAC	TTTTTGTCTTACGAGTGGGCGGA	1800	

```

Db 3774 AGAACATTCTCCTTAATCAATAGACATGCTATTGACTTTCTTACTCAACAAGATGGGGAGGA 3833
Qy 1801 ACATGCAAGGTGCTAGGACCTGATTGTTGCATAGGAATAGAAGATCTATCTTAAAAAATATC 1860
Db 3834 ACATGCAAGGTGCTTGGACCTGATTGTTGCATCGGATAGAGACTTGTTCAAAAAATATT 3893
Qy 1861 TCAGAACAAATCGACAAAATCAGAAAGGATGAACAAAGAGGAGAACTGGCTGGGGTCTA 1920
Db 3894 TCAGAGCAAAATGCACAAATTTAAAAAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG 3953
Qy 1921 GGTGGCAAAATGGTGGACATCTGACTGGGT 1950
Db 3954 GGTGGTAAATGGTGGACATCCGACTGGGT 3983

RESULT 6
US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; ; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 8256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Marburg GP (dTM)
US-10-491-121-32

Query Match 86.4%; Score 1767.6; DB 20; Length 8256;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 1 ATGAAGACCACATGTTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC 60
Db 1431 ATGAAGACCACATGTTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC 1490
Qy 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGTCGCGAACT 120
Db 1491 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGTCGCGAACT 1550
Qy 121 CTCAGAAAGACAGAAGACGTCCTCTGATCGGATTACACATGAGTGGGCAAAAAGTTGCT 180
Db 1551 CTCAGAAAGACAGAAGACGTCCTCTGATCGGATTACACATGAGTGGGCAAAAAGTTGCT 1610
Qy 181 GATTCCCTCTTTGAGGCGATCCAAAGCATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 1611 GATTCCCTCTTTGAGGCGATCCAGAGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 1670
Qy 241 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 300
Db 1671 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 1730
Qy 301 TCTGGAAAATCCTTGCTGTAGTATCCTCCTACCAACATCCGTGACTATCCGAAATGCAA 360
Db 1731 TCTGGAAAATCCTTGCTGTAGTATCCTCCTACCAACATCCGTGACTATCCTAAATGCAA 1790
Qy 361 ACTATCCATCATATTCAAGTCTCAAAACCCCTCATGCAACGGGGATCGCCCTTTCAATTATGG 420

```

Db 1791 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATG 1850
Qy 421 GGAGCATTTTTCTGTATGATCGCTTGCCTCCACAAAGTATGACGAGGCAAGTCTTC 480
Db 1851 GGGGCAATTTTTCTGTATGATCGCTTGCCTCCACAAAGTATGACGAGGCAAGTCTTC 1910
Qy 481 ACTGAAGGGAACATAGCAGCTATGATGTTGTCATAAGACAGTGCACAAAATGATTTTCTCG 540
Db 1911 ACTGAAGGGAACATAGCAGCTATGATGTTGTCATAAGACAGTGCACAAAATGATTTTCTCG 1970
Qy 541 CGGCAAGGACAGGATACCGTATATGATATCTGACTTCTACTTAATAATATGGAAGT 600
Db 1971 CGGCAAGGACAGGATACCGTATATGATATCTGACTTCTACTTAATAATATGGAAGT 2030
Qy 601 AGTAACGGAAACCAACGAATGACACTGGATGTTTCGGCGCTTCTCAAGAAATACAAATCT 660
Db 2031 AGTAACGGAAACCAACGAATGACACTGGATGTTTCGGCGCTTCTCAAGAAATACAAATCT 2090
Qy 661 ACAAGAAACCAACATGTGCTCCGTCCAAATATCTCCACCACTGCCACAGCCCGTCCG 720
Db 2091 ACAAGAAACCAACATGTGCTCCGTCCAAATATCTCCACCACTGCCACAGCCCGTCCG 2150
Qy 721 GAGATCAAACTCACAGCAACCCAACTGATGCGCAACAACTCAATACACGAGCCCAAGC 780
Db 2151 GAGATCAAACTCACAGCAACCCAACTGATGCGCAACAACTCAATACACGAGCCCAAGC 2210
Qy 781 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGTCGGAGACGAGAACCCACACA 840
Db 2211 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGTCGGAGACGAGAACCCACACA 2270
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 900
Db 2271 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 2330
Qy 901 CAACCAAGCAGCCGACAGCAAGGAGGAAACAACAACCAATTCCTCAAGATGCTGTGACT 960
Db 2331 CAACCAAGCAGCCGACAGCAAGGAGGAAACAACAACCAATTCCTCAAGATGCTGTGACT 2390
Qy 961 GAACCTAGACAAATAAACAACACTGCAACACCGTCCATGCCCCCTCATAACTACCAACA 1020
Db 2391 GAACCTAGACAAATAAACAACACTGCAACACCGTCCATGCCCCCTCATAACTACCAACA 2450
Qy 1021 ATCTCTACTAACACACACCTCCAAACACAACTTTCAGCACTCTCTGCGACCAATTAACAAAC 1080
Db 2451 ATCTCTACTAACACACACCTCCAAACACAACTTTCAGCACTCTCTGCGACCAATTAACAAAC 2510
Qy 1081 ACCCAATGACACACACAGACCAATCACTGAAATGAGCAAAACCAAGTGCCTCCCTCG 1140
Db 2511 ACCCAATGACACACACAGACCAATCACTGAAATGAGCAAAACCAAGTGCCTCCCTCG 2570
Qy 1141 ATAACAACCTCGCTCCAAACGGGAATCCACACAGCAAGAGGACCAAGCAAGCAAAAAA 1200
Db 2571 ATAACAACCTCGCTCCAAACGGGAATCCACACAGCAAGAGGACCAAGCAAGCAAAAAA 2630
Qy 1201 GGCCCGCCCAACACGCAACCAACGCAAAATGAGCAATTCACAGCTCCTCCCGCCAC 1260
Db 2631 GGCCCGCCCAACACGCAACCAACGCAAAATGAGCAATTCACAGCTCCTCCCGCCAC 2690
Qy 1261 CCAGCTCGACTGCAACAATCTGTATATTTTCAGAAAGCGATTCGATTTCTGGAAA 1320
Db 2691 CCAGCTCGACTGCAACAATCTGTATATTTTCAGAAAGCGAATATCTCTCGGAG 2750
Qy 1321 GAAGGTGATATATCCCGTTTTAGTGGTTTAAATAATGAAATGATTTTCATCCA 1380
Db 2751 GAAGGGGACATGTTCCTTTCTCGATGGTTTAAATAATGCTCCAAATGATTTTGACCCA 2810
Qy 1381 ATCCCAACACAGAAACCAATCTTGTATGAATCTCCAGCTTTTAATACTTCAACTAATGAG 1440
Db 2811 GTTCCAAATACAAACCAATCTTGTATGAATCTCTAGTTCTGGTGCCTCGCTGAGNA 2870
Qy 1441 GAACAAACACATCCCGCCGGAATATGATTTAACTTTCTCTTATTTTCTGTATTAATAATGA 1500

Db 2871 GATCAACATGCTCCCCAATATTTAGTTTAACTTTTCTTATTTTCTTAATAATAGAG 2930
Qy 1501 GATACTGCTACTCTGGGGAAACGAGAATGATTTGATGCAGAGTTTGAGGATTTGGAGT 1560
Db 2931 AACACTGCTACTCTGGAGAAATGAGAATGATTTGATGCAGAGTTAAGAAATTTGGAGC 2990
Qy 1561 GTGAGAGGAGCAGATTTGGCGGAGGCTTAGCTTAGCTAGCTACCATTTTGGCCCTGGAATC 1620
Db 2991 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAT 3050
Qy 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG 1680
Db 3051 GAAGGACTTTACACTGCTGTTTTAAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAG 3110
Qy 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTTTAAAGGGTCACAAACGAGAA 1740
Db 3111 CGCTAGCAATCAAACTGCTCAAAATCCTTGGAACTCTTTTATGAGAGTCACAACTGAGAA 3170
Qy 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGACTTTTGTCTACGAGTGGGGCGA 1800
Db 3171 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGAGGA 3230
Qy 1801 ACATGCAAGGTCAGGACCTGATTTGTCATAGGATAGAGATCTATCTAAAAAATATC 1860
Db 3231 ACATGCAAGGTCAGGACCTGATTTGTCATAGGATAGAGATCTTGTCCAAAATATTT 3290
Qy 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 3291 TCAGAGCAATTTGACCAAAATTAAGAGGACGAAACAAAGAGGGGACTGGTTGGGTCTG 3350
Qy 1921 GGTGGCAATGGTGACATCTGACTGGGGT 1950
Db 3351 GGTGGTAAATGGTGACATCTGACTGGGGT 3380

RESULT 7
US-10-811-353-2
; Sequence 2, Application US/10811353
; Publication No. US20050112098A1
; GENERAL INFORMATION:
; APPLICANT: McCray, Paul B.
; APPLICANT: Sanders, David A.
; APPLICANT: Jeffers, Scott A.
; APPLICANT: Davidson, Beverly L.
; APPLICANT: Sinn, Patrick L.
; TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE
; FILE REFERENCE: 290, 00670120
; CURRENT APPLICATION NUMBER: US/10/811,353
; CURRENT FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Marburg virus glycoprotein
US-10-811-353-2

Query Match 82.7%; Score 1692.4; DB 21; Length 2046;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
Qy 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 60
Qy 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAATATGGATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAATATGGATTCGGTATGCTCCGGAAT 120
Qy 121 CTCAGAGACAGAGACGTCCTATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAGACAGAGATGTCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180

181 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTCCAGGACAGGTGTACCTCCCAAGAAT 240
181 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTCCAGGACAGGTGTACCTCCCAAGAAT 240
241 GTTGAAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC 300
241 GTTGAAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC 300
301 TCTGGAATAATCTTGCTGTATGATCCTCTTACCAACATCCGTGACTATCCGAATGCAAA 360
301 TCTGGAATAATCTTGCTGTATGATCCTCTTACCAACATCCGTGACTATCCGAATGCAAA 360
361 ACTATCCATCATATTCAGAGTCAAAACCTCATGCGCAAGGGATCGCCCTCATTTGTGG 420
361 ACTATCCATCATATTCAGAGTCAAAACCTCATGCGCAAGGGATCGCCCTCATTTGTGG 420
421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACATGTATCCGAGGCAAAAGTCTTC 480
421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACATGTATCCGAGGCAAAAGTCTTC 480
481 ACTGAAGGAAACATAGCAGCTATGATTTCTCAATAGACAGTGCACAAATGATTTCTCG 540
481 ACTGAAGGAAACATAGCAGCTATGATTTCTCAATAGACAGTGCACAAATGATTTCTCG 540
541 CGGCAAGGACAGGATACCGTACATGATCTGACTTCTACTAATTAATATTTGACAGAT 600
541 CGGCAAGGACAGGATACCGTACATGATCTGACTTCTACTAATTAATATTTGACAGAT 600
601 AGTAACGGAACGCAAAATGACATGCTGATGTTTCGGCGCTCTTCAAGAAATACAATCT 660
601 AACATGGAACAAACCGATGACACTGATGCTTGGTGTCTTTCANGAATACACTCC 660
661 ACAAGAAACCAAAATGCTCGTCCGTCGCAAAATACCTCCACATGCCCACAGCCGCTCG 720
661 ACAAGAAACCAAAATGCTCGTCCGTCGCAAAATACCTCCACATGCCCACAGCCGCTCG 720
721 GAGATCAAACTCACAGGACACCCCACTGATGCCCAAACTCAATACACAGGCCCAAGC 780
721 GAGATCAAACTCACAGGACACCCCACTGATGCCCAAACTCAATACACAGGCCCAAGC 780
781 AGTGATGATGAGGACCTCGCAACATCGGGTCCGAGAACAGAACCCCAACA 840
781 AATGATGATGAGGACCTCATACATCCGGTTCAGGGTCCGAGAACAGAACCCCTATACA 840
841 ACTTCTGATGCGGTCAACAAGCAAGGGCTTTTCATCAAAATGCCACCACTCCCTCACCA 900
841 ACTTCTGATGCGGTCAACAAGCAAGGGCTTTTCATCAAAATGCCACCACTCCCTCACCA 900
901 CAACCAAGCACGCCACAGAGGAGGAAACAAACAAACCAATTCCTCCAAAGTGTGTGACT 960
901 CAACCAAGCACGCCACAGAGGAGGAAACAAACAAACCAATTCCTCCAAAGTGTGTGACT 960
961 GAACTAGACAAATAACACAACTGCACACCGTCCATGCCCTCATACACTACCA 1020
961 GAACTAGACAAATAACACAACTGCACACCGTCCATGCCCTCATACACTACCA 1020
1021 ATCTCTACTTAAACAACCTCCAAACACAACTTTCAGCACTCTCTGTCACCAATTAACAAAC 1080
1021 ATCTCTACTTAAACAACCTCCAAAGAACAACTTTCAGCACTCTCTGTCATCAACAAAC 1080
1081 ACCACAAATGACAAACACAGAGGACAAATCACTGAAATGAGCAAAACAGTGCCTCCCTCG 1140
1081 ACCACAAATGAGCAACACAGAGGACAGCCACTGAAATGAGCAAAACAGTGCCTCCCTCG 1140
1141 ATAACAACTCCCTCCCAAGGAAATCCCAACAGCAAAAGAGCAGCAGCAGCAAA 1200
1141 AAAACAACCTCCCTCCCAAGGAAATTCACAGCAAAAGAGCAGCAGCAGCAAA 1200
1201 GGGCCCGCCACAAACGCAACCAACGCAAAATGAGCAATTCACCAAGTCCCTCCCTCCAC 1260
1201 GGGCCCGCCACAAACGCAACCAATATGACAAATGGGCAATTTAACCAAGTCCCTCCCTCC 1260
1261 CCCAGCTCGACTGCAACAATCTTGATATTTTTCAGAGAAAGCGGATCGATTTCTCGGAAA 1320

1261 CCCAACCCGACCAACAACATCTGTATATTTTCAAGAAAGAACGAGTATCTCTCGAGG 1320
1321 GAAGGTGATATATTTCCGTTTTTAGATGGGTAAATAAATACTGAAATTTGATTTGATCCA 1380
1321 GAAGGCGACATGTTCTCTTCTGGACGGGTAAATAAATGCTCCAAATGATTTGATTTGATCCA 1380
1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCGAGCTTTAATATCTCAACTAATAGAG 1440
1381 GTTCCAAATCAAGACGATCTTTGATGAATCTTCTAGTTCTGGTCTTCGGCTCGAGGAA 1440
1441 GAACAACACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAAAATGGA 1500
1441 GATCAACATGCTCTCCCAATATCAGTTTAACTTTATCTTATTTCTTAATATAAATGAA 1500
1501 GATACTCCCTACTCTGGGGAACGAGAATGATTTGATGCAGAGTTTGAGGATTTGGAGT 1560
1501 AACACTGCTACTCTGGAGAAATGAGAACGATTTGTGATGCAGAGTTTAAGAAATTTGGAGC 1560
1561 GTGAGAGGACAGATTTGGCGGACGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
1561 GTTCAGGAGGATGACCTGGCAGCAGGGCTCAGTTGGATACCGTTTCTTGGCCCTGGAATC 1620
1621 GAAGACTCTATACCTGCGGTTTAAATCAAAATCAGAACAAATTTAGTTTGTAGTTGAGG 1680
1621 GAAGACTTTTATCTGCTGTTTAAATTAATAAACCAAAACAAATTTGGTCTGCAGGTTGAGG 1680
1681 CGCTTACTTAATCAAACTGCTTAAATCCTTGGAGCTCTTGTAAAGGGTCACAAACGAGGAA 1740
1681 CGTCTAGCAATCAAACTGCAAAATCCTTGGAACTCTTATTAGAGTCACAAACGAGGAA 1740
1741 AGGACATTTCTTAAATAGAGCATGCAATTTGACTTTTGTCTTACAGAGTGGGGCGGA 1800
1741 AGGACATTTCTTAAATTAATAGACATGCCAATTTGACTTTCTACTCACAAGGTGGGGAGGA 1800
1801 ACATGCAAGGTGCTAGGACCTGATTGTTGCATAGAGATAGAAGATCTATCTTAAATAATATC 1860
1801 ACATGCAAGGTGCTTGGACCTGATTGTTGCATTTGAAATAGAAGACTTGTCCAGGAATATT 1860
1861 TCAGAACAAATTCGACAAATCAGAAAGGATGACAAAGGAGGAAACTGGCTGGGGTCTA 1920
1861 TCGAACAAATTCAGCAAAATCAAAAGAGATGACAAAGGAGGACTGGTTGGGGTCTA 1920
1921 GGTGGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCATCTCTGCTACTA 1980
1921 GGTGGTAAATGGTGACATCCGACTGGGGTGTCTTACTAACTTTGGGCATTTTGGCTACTA 1980
1981 TTATCTATAGTGTCTGATTGCTCTGCTGCTGATCTGCTGATCTTCACTAATATCAATT 2040
1981 TTATCCATAGTGTCTGATTGCTCTATCTCTGCTGATTTGTGCTATCTTTTACCAAAATATATC 2040
2041 GGATGA 2046
2041 GGGTAA 2046

RESULT 8

US-10-353-856-27

; Sequence 27, Application US/10353856

; Publication No. US20030215794A1

; GENERAL INFORMATION:

; APPLICANT: Kawaoka, Yoshihiro

; APPLICANT: Jasenosky, Luke D.

; APPLICANT: Neumann, Gabriele

; APPLICANT: Wisconsin Alumni Research Foundation

; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Base

; FILE REFERENCE: 800.032US1

; CURRENT APPLICATION NUMBER: US/10/353,856

; PRIOR FILING DATE: 2003-01-29

; PRIOR APPLICATION NUMBER: US 60/353,972

; PRIOR FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSEQ for Windows Version 4.0


```
; SEQ ID NO 27
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Marburg virus
US-10-353-856-27

Query Match      82.7%; Score 1692.4; DB 17; Length 19112;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTCTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCC 60
Db 5940 ATGAAGACACATGTTCTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCC 5999

QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059

QY 121 CTCGAGAAGACAGAAGAGCTTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCGAGAAGACAGAAGAGATGCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119

QY 181 GATTCCTCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 6120 GATTCCTCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 6179

QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC 300
Db 6180 GTTGAGTATACAGAGGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC 6239

QY 301 TCTGAAAATCTCTGTTTAGATCCTCTACCAACATCCCGTGAATCTCCGAAATGCAAA 360
Db 6240 TCTGAAAATCTCTGTTTAGATCCTCTACCAACATCCCGTGAATCTCCGAAATGCAAA 6299

QY 361 ACTATCCATCATATTCNAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTCNAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG 6359

QY 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC 480
Db 6360 GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC 6419

QY 481 ACTGAAGGAACATAGCAGCTATGATGTCAATAAGACAGATGCAACAAATGATTTTCG 540
Db 6420 ACTGAAGGAACATAGCAGCTATGATGTGTCAATAAGACAGATGCAACAAATGATTTTCG 6479

QY 541 CGGCAAGNACRAGGCTACCGTCAATGATCTGACTTCTACTAATAATATGGCAAGT 600
Db 6480 AGGCAAGGACAGGGGTACCGTCAATGATCTGACTTCTACTAATAATATGGCAAGT 6539

QY 601 AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATTC 660
Db 6540 AACATGGNACACAAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATTC 6599

QY 661 ACAAGAACCAAAATGCTCGTCCGTCGAAATACCTCCACCACTGCGCCACAGCCCGTCG 720
Db 6600 AGCAAGAAATCAAAATGCTCGTCCGTCGAAATACCTCCACCACTGCGCCACAGCCCGTCG 6659

QY 721 GAGATCMAACTCACAGACACCCCACTGATGCCCACTCAATACCAAGCCCAAGC 780
Db 6660 GAGATCMAACCCCAAGACACCCCACTGATGCCCACTCAATACCAAGCCCAAGC 6719

QY 781 AGTGATGTAGGACTCTCGCAACATCTCGGCTCAGGGTCGGAGAACGAGAACCCCAACA 840
Db 6720 ANTGATGTAGGACTCTCTATACATCTCGGTTTCAGGGTCGGAGAACGAGAACCCCTATACA 6779

QY 841 ACTTCTGATGCGGTCAACGAAGAGGGCTTTCATCAACAAATGCCACCACTCCCTCACA 900
Db 6780 ACTTCAAGTGCCTCACTAAGCAAGGGCTTTCATCAACAAATGCCACCACTCCCTCACA 6839

QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCACAGCAAGGAGGAAACACACAAACCATTTCCCAAGGTAATGCTGTGACT 6899
```

```
QY 961 GAACCTAGACAAAAATAACAACTGCAACAACCGTCCATGCCCCCTCTATAACACTACCACA 1020
Db 6900 GAACCCAAACAAAACCAACACAAACGGCAACAACCGTCCATGCCCCCACCACAAACCACTGCA 6959

QY 1021 ATCTCTACTATAACAACACTCCCAACAACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
Db 6960 ATCTCTACTATAACAACACTCCCAAGAACAACACTTTCAGCACTCTCTCTGTATCTACAAAC 7019

QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAATAGCAAAACAGTCCGCCCTCG 1140
Db 7020 ACCACCAATGACACACAGAGCACAAGCCACTGAAATAGCAAAACAGTCCGCCCTCG 7079

QY 1141 ATAAACAACCTCTCCCAACGGGAAATCCCAACAGCAAAAGAGCAGCAGCAAGAAAA 1200
Db 7080 AAAACAACCTCTCCCAACAGGAAATCTTACCACAGCAAAAGAGCAGCTAACACACGAAA 7139

QY 1201 GGCCCGGCACAAACGGGACCAACACAGCAAAATGAGCAATTCACAGTCTCCGCCAC 1260
Db 7140 GGCCCGGCACCAACGGGACCAAAATATGCAAAATGGCATTTAACAGTCTCCGCCAC 7199

QY 1261 CCCAGCTCGACTGCAACAACATCTTTGTATATTTCAAGAAAGAGCGATCGATTTCTGAAA 1320
Db 7200 CCCAACCGGACCAACAACATCTTTGTATATTTCAAGAAAGAACGAGTATCTCTGGAGG 7259

QY 1321 GAAGGTGATATATTTCCGTTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Db 7260 GAAGGCGACATGTTTCTCTGACGGGTTAATAATGCTCCAAATTTGATTTGATCCA 7319

QY 1381 ATCCCAACACAGAAACAACTTTTGTATGATCTCCAGCTTTAATACTCAACTAATGAG 1440
Db 7320 GTTCCAAATACAAAGACGATCTTTGATGAAATCTTCTAGTTCTGGTGTCTCGGCTGAGAA 7379

QY 1441 GAAACAACACACTCCCGCAATATCAGTTTAACTTTTCTTATTTTCTTGATAAAATGGA 1500
Db 7380 GATCAACATGCTCCCCCAATATCAGTTTAACTTTTCTTATTTCTTAATAATAAGAA 7439

QY 1501 GATACCTCTACTCTGGGGAAACGAGAAATGATTTGATGAGAGTTGAGGATTTGGAGT 1560
Db 7440 AACACTGCTACTCTGGAGAAATGAGAACGATTTGTGATGAGAGTTAAGAAATTTGGAGC 7499

QY 1561 GTGAGAGGAGACGATTTGGCGGAGGGCTTAGCTGGATACCAATTTTGGCGCTCGGAATC 1620
Db 7500 GTTCAGGAGGATGACCTGGCAGAGGGCTCAGTTGGATACCGTTTTTGGCGCTCGGAATC 7559

QY 1621 GAAGACTCTATACTGCGGTTTTAATCAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG 1680
Db 7560 GAAGACTTTTATCTGCTGGTTTTAATAAAACCAAAACAATTTGGTCTGCAGGTTGAGG 7619

QY 1681 CGCTTAGCTAATCAAACTGCTAAAATCCTTGGAGCTCTTGTAAAGGTCACAAACGAGAA 1740
Db 7620 CGCTTAGCCAATCAAACTGCTAAAATCCTTGGAACTCTTATTAAGAGTCAACAACGAGAA 7679

QY 1741 AGGACATTTTCTTAATCAATAGCATGCAATTCACCTTTTGGCTTACGAGGTGGGCGGA 1800
Db 7680 AGGACATTTTCTTAATTAATAGACATGCCATTCACCTTCTACTCACAAGGTGGGAGGA 7739

QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTGATAGAAATAGAAATCTATCTATAAAATATC 1860
Db 7740 ACATGCAAGGTGCTTGGACCTGATTTGTGATAGAAATAGAAATCTTGTCCAGGNATAT 7799

QY 1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAACAAAGAGGAGAACTGGCTGGGGTCTA 1920
Db 7800 TCGGAACAAATTCGACAAATCAGAAAGATGAACAAAGAGGAGGACTGGTTGGGGTCTA 7859

QY 1921 GGTGGCAATGCTGGACATCTGACTGGGTGTTCTCACCATTTGGGGATCTCTGCTACTA 1980
Db 7860 GGTGGTAAATGCTGGACATCTGACTGGGGTGTCTTACTAACTTTGGGATTTTGTACTA 7919

QY 1981 TTATCTATAGTGTCTTGATTTGCTCTGTCCTGTATCTCTGCTATCTTCACTAAATACAT 2040
Db 7920 TTATCCATAGTGTCTTGATTTGCTCTATCCGTATTTGCTGTATCTTTCACCAATATATC 7979
```


Qy 2041 GGATGA 2046
Db 7980 GGGTAA 7985

RESULT 9

US-10-066-506A-13
; Sequence 13, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan. L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 13
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Marburg virus strain Raven
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-13

Query Match 74.2%; Score 1518; DB 15; Length 2046;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 1716; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

Qy 1 ATGAAGACACATGTTTCCTTATAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 60

Qy 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAAAATTGCTAGTAACAGCCAACTCAAGATGTAGATTCAAGTGTGCTCCGGAAC 120

Qy 121 CTCGAGAGACAGAGAGCTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGAGACAGAGAGTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180

Qy 181 GATTCCTCTTTCGAGGCAATCAAGCGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCTCTTTCGAGGCAATCAAGCGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT 240

Qy 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC 300
Db 241 GTTGAGTATACGAGGAGAGAGCCAAACATGTTACAATATAAGTGTAAACAGCCCT 300

Qy 301 TCTGAAATCTTGTGTTAGTATCTCTCAACATCCGTTGATATCCGAAATGCGAAA 360
Db 301 TCTGAAATCTTGTGTTAGTATCTCTCAACATCCGTTGATATCCGAAATGCGAAA 360

Qy 361 ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG 420
Db 361 ACTGTTTCATATATCAAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG 420

Qy 421 CGAGCATTTTCTGTATATCGATGCTCCCAACAAATGTACCGAGGCAAGTCTTC 480
Db 421 GGGGCATTTTCTGTATATCGATGCTCCCAACAAATGTACCGAGGCAAGTCTTC 480

Qy 481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATGTCAATAGACAGTTCACAGAATGATTTTCT 540

Qy 541 CGGCAAGGCAAGGGTACCGTATATGATCTGACTTCTACTAATAATATTTGGACAGT 600
Db 541 AGGCAAGGCAAGGGTATCTGCATCAATGAATTTGACCTCCACCANTAAATATTTGGCAAGC 600

Qy 601 AGTAAACGGAACGCAACGAATGACACATGGATGTTTTCGGCGCTCTTCAAGAATACAAATCT 660
Db 601 AGCAATGAACGCGAGAGAAATGATACGGGATGTTTTCGGCATCTCTCCAGAAATACAACTCC 660

Qy 661 ACAAGAACCAAAACATGTGCTCCGTCGCAAAATACCTCCACCACATGCCCCACAGCCGCTCG 720
Db 661 ACAAAATCAAAATCATGCTCCATCTCTTAACTCCATCTCCCTGCCCCACAGTAACTCG 720

Qy 721 GAGATCAAACTCACAGCACCCCAACTGATGCCCAAACTCAATACACAGGACCAAGC 780
Db 721 AGCATTCACCTACAAATACCTCAAAATTAATACTGCTAAATCTGGAACCTATGAACCAAGT 780

Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAACGAGAACCCCAACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840

Qy 841 ACTTCTGATGGGTCCCAAGCAAGGGCTTTCATCAACAAATGCCACCCACCTCCCTCACA 900
Db 841 ACTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

Qy 901 CAACCAAGCACGCCACAGCAAGGAGGAAACAAACAAACCAATTCACCAAGATGCTGTGACT 960
Db 901 CATCCAGCACCTCACAACATGAGCAAAACAGTAGAATCTCTCCGACATGCTGTAAT 960

Qy 961 GAATAGACAAAATAACAACTGCGACAAACCGTCCATGCGCCCTCATATAACACTACACA 1020
Db 961 GAGCACAAATGGAACCGACCCCAACCAACCAACAGCAACGCTCCTCAACAAATCTAATA 1020

Qy 1021 ATCTCTACTTAAACACCTCCAAACACAACTTTCAGCACTCTCTGTCACCACTTACAAAAC 1080
Db 1021 ACTCCACCTTAAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080

Qy 1081 ACCACCAATGACAAACACAGAGACCAATCACTGAAAATGAGCAAAACAGATGCCCTCG 1140
Db 1081 ATCCAAATATGATACACAGCTGAACTAGCAGAAAGCAACAAACCAATGCTCAGTTG 1140

Qy 1141 ATAAACACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGACACAGAGGCAAAA 1200
Db 1141 AACCAACTCTAGATCCAAACAGAAAATCCCAACAGCAAAAGACACCAACAGCAACAAC 1200

Qy 1201 GGCCCCGACACAGGCAACCAACAGCAAAATGAGCAATTTCCAGCTCTCCCTCCACCC 1260
Db 1201 AACATCATGACGACATCAGATATACAAAGCAAAACCCCAACAAATCTCTCCGGAT 1260

Qy 1261 CCCAGCTCGATGCAACCAATCTTGTATATTTAGAGAAAGAGCGATCGATTTCTGGAAA 1320
Db 1261 TCTAGTCCGACAAACCGCCCTCTTATATCTTTAGAAAGAAACGATCGATTTCTGGAAA 1320

Qy 1321 GAAAGTGATATATTTCCGTTTTTATAGTGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Db 1321 GAAAGTGATATATTTCCGTTTTTATAGTGGTTAATAATACTGAAATTTGATTTGATCCA 1380

Qy 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTTAACTTCAACTAATGAG 1440
Db 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTTAACTTCAACTAATGAG 1440

Qy 1441 GAAACAAACACTCCCGCCGAATATCAGTTTAACTTTCTTATTTCTCTGATATAAAATGGA 1500
Db 1441 GAAACAAACACTCCCGCCGAATATCAGTTTAACTTTCTTATTTCTCTGATATAAAATGGA 1500

Qy 1501 GATCTGCTACTCTGGGGAAAAACGAGAAATGATGTGATGACAGATTTGAGGATTTGGAGT 1560
Db 1501 GATCTGCTACTCTGGGGAAAAACGAGAAATGATGTGATGACAGATTTGAGGATTTGGAGT 1560

Qy 1561 GTGAGGAGGAGGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
Db 1561 GTGAGGAGGAGGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620

Qy 1621 GAAGGACTCTATCTGCGGGTTTTTAATCAAAAATCAGAAACAAATTTAGTTTGTAGGTTGAGG 1680
Db 1621 GAAGGACTCTATCTGCGGGTTTTTAATCAAAAATCAGAAACAAATTTAGTTTGTAGGTTGAGG 1680

```
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCCTTGGAGCTCTTGTAAAGGGTCACAAACGAGGAA 1740
Db |||||
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCCTTGGAGCTCTTGTAAAGGGTCACAAACGAGGAA 1740
Db |||||
QY 1741 AGGACATTTTCTTAATCAATAGGCATGCAATGACTTTTTCCTTACGAGGTGGGGCGGA 1800
Db |||||
QY 1741 AGGACATTTTCTTAATCAATAGGCATGCAATGACTTTTTCCTTACGAGGTGGGGCGGA 1800
Db |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAAAAATATC 1860
Db |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAATCTATCTAAAAATATC 1860
Db |||||
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
Db |||||
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
Db |||||
QY 1921 GGTGGCAATGTGGACATCTGACTGGGGTGTCTCACCNAATTTGGGCATCTGCTACTA 1980
Db |||||
QY 1921 GGTGGCAATGTGGACATCTGACTGGGGTGTCTCACCNAATTTGGGCATCTGCTACTA 1980
Db |||||
QY 1981 TTATCTATAGCTGTCTGATTCCTGCTCTGCTATCTGTCTGATCTTCACTAAATACATT 2040
Db |||||
QY 1981 TTATCTATAGCTGTCTGATTCCTGCTCTGCTATCTGTCTGATCTTCACTAAATACATT 2040
Db |||||
QY 2041 GGATGA 2046
Db |||||
QY 2041 GGATGA 2046
Db |||||
```

RESULT 10

```
US-10-066-506a-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein
; OTHER INFORMATION: Marburg virus strain Musoke Glycoprotein 2
US-10-066-506a-7
```

```
Query Match 64.5%; Score 1319.6; DB 15; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;
```

```
QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGGACAAAAATCTCCCC 60
Db |||||
QY 1 ATGAAGACACATATATTTCTGATAGTCTCTCAATTTAATCCAAAGTATAAAAACTCTCCCT 60
Db |||||
QY 61 ATTTTAGACATAGCTAGTAATATCAACCCCAAAATGTGATTCGGTATGCTCCGGAACT 120
Db |||||
QY 61 GTTTTAGAAATTTGCTAGTAACGCCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACT 120
Db |||||
QY 121 CTCGAGAGACAGAAAGACGTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db |||||
QY 121 CTCGAGAGACAGAAAGATGTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db |||||
QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db |||||
QY 181 GATTCCTCTTTGGAGGCATCTAAACGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db |||||
```

```
QY 241 GTTGAGTACACAGAGGGGAGGAAAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 300
Db |||||
QY 241 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTACAATAAAGTGTAAACGACCT 300
Db |||||
QY 301 TCTGAAAAATCCTTGCTGTGTAGATCCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db |||||
QY 301 TCTGAAAAATCCTTGCTGTGTAGATCCCTCTACCAACATCCGTGACTATCCCTTAAATGTAAA 360
Db |||||
QY 361 ACTATCGATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db |||||
QY 361 ACTGTTCAATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATTCCTCTTCATTTGTTG 420
Db |||||
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCACAACAATGTATCCGAGGCAAAAGTCTTC 480
Db |||||
QY 421 GGAGCATTTTCTGTATGATCGCGTTGCTCTCAACAACAATGTATCCGAGGCAAAAGTCTTC 480
Db |||||
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db |||||
QY 481 ACTGAAGGGAATATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db |||||
QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 600
Db |||||
QY 541 AGGCAAGGACAAAGGTATTCGTCAATGAACCTTGACCTCCACAATAAATATTTGGACAAGC 600
Db |||||
QY 601 AGTAAACGGAACGCAAAACGAATGACATGCGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 660
Db |||||
QY 601 AGCAATGAACGCGAGAGAAATGATACGGGATGTTTTCGGCATCTCCCAAGAATACAACTCC 660
Db |||||
QY 661 ACAAAGAACCAAAACATGTGCTCCGTCCAAATATCTCCACCATGCCCCACAGCCCGTCCG 720
Db |||||
QY 661 ACAAACCAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
Db |||||
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGTCGCAAACTCAATACCAACGACCCCAAGC 780
Db |||||
QY 721 AGCATTTCACTCTCAAAATACATCAAAATTAATATCTGTAATCTGGAACTATGAACCAAGT 780
Db |||||
QY 781 AGTGATGATGAGGACCTCGCAACATCTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACACA 840
Db |||||
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db |||||
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db |||||
QY 841 ACTCTTAATGTAGTCACTGAAACAGAAAACAATCGTCAACAATATTTGTCTCACTCTTCACTA 900
Db |||||
QY 901 CAACCAAGCAGCCACAGAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT 960
Db |||||
QY 901 CATCCAAGCACCTCACAACTGAGCAAAACAGTAGAATCTCTCCGACATGCTGTAACT 960
Db |||||
QY 961 GAACTAGACAAAAATAACAACTGCAACACCGTCCATGCCCTCATGCCCTCATAAACACTACCACA 1020
Db |||||
QY 961 GAGCACAATGGAAACCGACCCCAACAAACAAACCAACGCAACGCTCTCAACAATACTAATA 1020
Db |||||
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTCTGACCACTTACCAATTA 1080
Db |||||
QY 1021 ACTCCCACTTAACACTCTCAAGTACAACTCAGTACTCTCTTCCCTTCCCAACCCGCAAC 1080
Db |||||
QY 1081 ACCACCAATGACACACAGACACATCACTGAAATGAGCAAAACCAAGTGCCTCCCTCG 1140
Db |||||
QY 1081 ATCAACCAATATGATACACAACGCTGAACCTAGAGAAAGCGAACCAACCAATGCTCAGTTG 1140
Db |||||
QY 1141 ATAACCAACCTTCCTCCAAACGGGAAATCCCAACGACAGCAAAAGAGCACAGGACGCAAAAA 1200
Db |||||
QY 1141 AACCAACTCTAGATCCAAACAGAAAAATCCCAACGACAGCAAGACACCAACAGCAACAAC 1200
Db |||||
QY 1201 GGCCCCGCCACAAAGGCAACCAACAGCAAAATGAGCAATTCACGATCTCTCCCTCCAC 1260
Db |||||
QY 1201 AACATCATCATGAGCATCAGATATACAGAGCAAAACCCCAACAAATTTCTTCTCCGAT 1260
Db |||||
QY 1261 CCCAGCTCGCTGACACACATCTTGTATATTTTCAAGAAAGGAGGATCGATTTCTTGGAAA 1320
Db |||||
QY 1261 TCTAGTCCGACAAACCCGCCCTCTTATATCTTTAGAAAGAACGATCGATCTCTTGGAGG 1320
Db |||||
```

```
QY 1321 GAAGGTGATATATCCCGTTTTAGATGGTTAATAAATACAGAAATGATTTTGATCCA 1380
Db 1321 GAAGCGACATGTTCCCTTTCTGTGATGGTTAATAAATGCTCCAATGATTTTGACCCA 1380
QY 1381 ATCCAAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCTGTGGTCCCTCGGCTGAGAA 1440
QY 1441 GAAACAACACATCCCGCCGAATACAGTTTAACCTTTCTTATTTCTTGTGATAAATAATGGA 1500
Db 1441 GATCAACATGCTCCCGCCGAATACAGTTTAACCTTTCTTATTTCTTGTGATAAATAATGAG 1500
QY 1501 GATAGTCTACTCTGCGGGAACACAGAAATGATTTGTGATGAGAGTTCAGAGATTTGGAGT 1560
Db 1501 AACACTGCTACTCTGCGGGAACACAGAAATGATTTGTGATGAGAGTTCAGAGATTTGGAGC 1560
QY 1561 GTGAGGAGGAGCAATTTGCGCGCAGGCTTAGCTGGATACCAATTTTTTTGGCCCTGGAAATC 1620
Db 1561 GTTCAGGAGGATGACCTGCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAAT 1620
QY 1621 GAAGGACTTATACCTGCGGTTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680
Db 1621 GAAGGACTTATACCTGCTGTTTTAATTAATAAATCAAAAACAATTTGGTCTGCAAGTTGAGG 1680
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTTGTTAAGGTCACAAACCGAGAA 1740
Db 1681 CGCTTAGCAATCAAACTGCGCAATCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1740
QY 1741 AGGACATTTCTTAATCAATAGGATGCAATTTGACTTTTTTTCCTTACGAGTGGGGCGGA 1800
Db 1741 AGAACATTTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGAGGA 1800
QY 1801 ACATGCAAGTGTCTAGGACCTGATTGTTGCATAGAAATAGAAATCTATCTAAAAAATATC 1860
Db 1801 ACATGCAAAAGTGTCTGACCTGATTGTTGCATCGGGATAGAAATCTGTCCAAAAATAT 1860
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGAGGAGGAACCTGGCTGGGCTCTA 1920
Db 1861 TCAGAGCAAAATGACCAAAATGAAAGAGGACGACAAAGAGGAGGAGTGGTTGGGGCTG 1920
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTCTCACCAAATTTGGGCAATCTGCTACTA 1980
Db 1921 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTTAATCTGGGCAATTTGCTACTA 1980
QY 1981 TTATCTAGTGTCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 TTATCCATAGTGTCTGATGCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 2041 GGATGA 2046
Db 2041 GGATAA 2046
```

RESULT 11

```
US-10-066-506A-3
; Sequence 3, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506A-3
```

```
Query Match          63.9%; Score 1308; DB 15; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
QY 121 CTCAGAAAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAAAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
Db 181 GATTCCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
QY 241 GTTAGGTACACAGAGGGGGAGGAGCCAAACATGCTACAATTAAGTGTAAAGGATCCC 300
Db 241 GTTAGGTACACAGAGGGGGAGGAGCCAAACATGCTACAATTAAGTGTAAAGGATCCC 300
QY 301 TCTCGAAAAATCTCTTGTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTCGAAAAATCTCTTGTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGAGCAAGTCTTC 480
Db 421 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGAGCAAGTCTTC 480
QY 481 ACTCAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
Db 481 ACTCAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
QY 541 CGGCAAGGACAAAGGATACCGTCTATGAAATCTGACTTCTACTAATAAATATTTGGACAAGT 600
Db 541 CGGCAAGGACAAAGGATACCGTCTATGAAATCTGACTTCTACTAATAAATATTTGGACAAGT 600
QY 601 AGTAAACGGAACGCAAAACGAATGACATGAGTGTTCGGGCTCTTCAAGAAATACAATTTCT 660
Db 601 AGTAAACGGAACGCAAAACGAATGACATGAGTGTTCGGGCTCTTCAAGAAATACAATTTCT 660
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATATCTCCACACTGCCACACAGCCGCTCG 720
Db 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATATCTCCACACTGCCACACAGCCGCTCG 720
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
Db 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
QY 841 ACTTCTGATGCGGTCAACAGCAGGAGGCTTTCATCAAAATGCCACCTCCCTCACA 900
Db 841 ACTTCTGATGCGGTCAACAGCAGGAGGCTTTCATCAAAATGCCACCTCCCTCACA 900
QY 901 CAACCAAGCACGCCACAGCAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
Db 901 CAACCAAGCACGCCACAGCAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
```

QY 961 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCCCTCATAAACAATACCAACA 1020
DB 961 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCCCTCATAAACAATACCAACA 1020
QY 1021 ATCTCTACTTAAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCAACCAATACAAAAC 1080
DB 1021 ATCTCTACTTAAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCAACCAATACAAAAC 1080
QY 1081 ACCACCAATGACAAACACACAGACGACAACTCACTGAAAATGAGCAAAACAGTCCCTCCG 1140
DB 1081 ACCACCAATGACAAACACACAGACGACAACTCACTGAAAATGAGCAAAACAGTCCCTCCG 1140
QY 1141 ATAAACAACCTCCCTCCAAACCGGAAATCCACCAAGAGAGACCAAGAGCAAAACAA 1200
DB 1141 ATAAACAACCTCCCTCCAAACCGGAAATCCACCAAGAGAGACCAAGAGCAAAACAA 1200
QY 1201 GGGCCGGCCACAAACCGGCAACAGCAACCAATGAGCATTTTACCAGTCTCTCCCTCCG 1260
DB 1201 GGGCCGGCCACAAACCGGCAACAGCAACCAATGAGCATTTTACCAGTCTCTCCCTCCG 1260
QY 1261 CCAGCTGCACTGCAACCACTCTGTATATTTTCAAGAAAGCGATCG 1308
DB 1261 CCAGCTGCACTGCAACCACTCTGTATATTTTCAAGAAAGCGATCG 1308

RESULT 12

US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 6902
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized)
US-10-491-121-42

Query Match 48.7%; Score 995.6; DB 20; Length 6902;
Best Local Similarity 69.5%; Pred. No. 1.1e-284;
Matches 1352; Conservative 0; Mismatches 594; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTTCCCTTATCATGCTTATCTTAAATTCAGGGGCAAAAATCTCCCC 60
DB 1923 ATGAAGACCAATGTTTCCCTTATCATGCTTATCTTAAATTCAGGGGCAAAAATCTCCCC 1982
QY 61 ATTTTATAGATAGCTAGTAAATCAACCCCAAAATGTGATTCGGTATGCTTCCGGAAT 120
DB 1983 ATCTTGGAGTCGCCAGCAACACAGCCCAAGACGTGGAGCGTGTGACGGCACC 2042
QY 121 CTCGAGAGACAGACGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGTTCT 180
DB 2043 CTCGAGAGACGAGACGCTCCATCTGATGGGCTTCACTTACCGGGCCAGAGGTGCC 2102
QY 181 GATTCCTCTTGGAGCATCCAGCGATGGGCTTTCAGACAGGTGTACCTCCCAAGAT 240
DB 2103 GACAGCCCTCTGGAGGCCAGCAAGAGGTGGGCTTTCAGAGCCGGCGTGGCCCAAGAAC 2162

QY 241 GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACAAATATAAGTGTAAACGATCCC 300
DB 2163 GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACAAATATAAGTGTAAACGATCCC 2222
QY 301 TCTGGAAAATCCTTGTCTTATAGATCCCTTACCAACATCCGCTGACTATCCGAAAATGCAAA 360
DB 2223 AGCGCAAGAGCCCTGTCTGTGACCCCTCCCAACCAATCATCGGGACTACCTTAAGTGCAG 2282
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
DB 2283 ACCATCCACCATCTCCAGGGCCAGAACCTTCACGCCAGGGCATCGCCCTGCACCTGTGG 2342
QY 421 GGAGCATTTTTTCTGTATGATGCTGCTCCCAACAATGTATCCGAGGCAAAATGCTTTC 480
DB 2343 GGGCGCTTCTTCTGTATGACAGAGGATCGCCAGCACCACTGTATACAGGGGAGGGGTTC 2402
QY 481 ACTGAAGGACATAGCAGCTATGATTTGTCAATGACAGTGCACAAAATGATTTTCTCG 540
DB 2403 ACCGAGGCAACATCGCGCCATGATCGTTAAACAGACCGGTGCACAAAGATGATCTTCAGC 2462
QY 541 CGGCAAGGACAAAGGTACCGTCAATATGATCTGACTTCTTACTAATAAATATTTGGACAAAGT 600
DB 2463 AGGCAAGGCGAGGGCTACAGGCACATGAACTGACCCAGCAACCAAGTACTTGGACCCAGC 2522
QY 601 AGTAAACGAAACGCAAAACGAATGACATGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 660
DB 2523 AACAAACGCAACAGACCAACGACACCGCTGCTTCCGCGCCCTGACAGAGTACAAACAGC 2582
QY 661 ACAAGAACCAAAACATGTGCTCCGTCGCAAAATACCTCCACACATGCCACACAGCCCGTCCG 720
DB 2583 ACCAAGAACCCAGACCTCGCGCCCGCAGCAAGATGCCAGCCCTGCCCACCGCCAGGGCCC 2642
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCGACCAAACTCAATACCAAGACCCCAAGC 780
DB 2643 GAGATCAAGCCCAACAGCACCCCAACCGACGCGCACCACTTGNACACCAACCCGACCCCAAC 2702
QY 781 AGTATGATGAGGACCTCGCAACATCGCGCTCAGGGTCCGAGAAACGAGAAACCCCAACA 840
DB 2703 AACGACGAGGAGGACCTGATCACAGCGGACGCGGCGAGGAGGAGGAGGAGGAGGAGGAGG 2762
QY 841 ACTTCTGATGGGTACCAACAGAGGGCTTTTCATCAACATGCCACCCACTCCCTCAACA 900
DB 2763 ACCAGCGACGCGGTGACCAACAGCGGGCTGAGCAGCAGCACCCTGCTTACCCCTAGCCCT 2822
QY 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACCAAACTTCCCAAGATGCTGTGACT 960
DB 2823 CAGCCAGCACCCCTCAGCAGAGGGCAACAAACCGACACACGCGGACGCGGACCGGACCGG 2882
QY 961 GAACCTAGACAAAATAACAACTGCAACACCGTCCATGCCCCCTCATAAACAATACCAACA 1020
DB 2883 GAGCCCAACAAGACCAACACACCGCCCAAGCATGCTCTCTCAACAACCAACCCGCGCC 2942
QY 1021 ATCTCTACTTAAACAACCTCCAAACACAACTTTCAGCACTCTCTGCAACCAATACCAAAAC 1080
DB 2943 ATCAGCACCAACACACAGCAAGAAACAACTTTCAGCACTCTCTGAGGCTGAGGCTGAGGAG 3002
QY 1081 ACCACCAATGACAAACACACAGACGACAACTCACTGAAAATGAGCAAAACAGTCCCTCCG 1140
DB 3003 ACCACCAATGACAAACACAGCAAGCAACCGCCACCGGACGAGCAGCAGCAGCAGCAGCAGC 3062
QY 1141 ATAAACAACCTCCCTCCAAACCGGAAATCCCAACAGCAAGCAAGAGAGACCAAGAGCAAAAC 1200
DB 3063 AAGACCACTTCTCCCAACCGGCAACCTGACCAACCGGCAAGAGAGAGCAACCAACCAACCAAG 3122
QY 1201 GGGCCGGCCACAAACCGGCAACCAAGCAAAATGAGCATTTTACCAGTCTCTCCCTCCG 1260
DB 3123 GGGCCCAACCAACCGGCTTAAACATGACCAACCGGCCACCTGACAGCCCGCCAGCCCGCCAGC 3182
QY 1261 CCCAGCTGCACTGCAACCAATCTTGTATATTTTCAAGAAAGCGATCGATTTTCTGGAAA 1320
DB 3183 CCCAACCCCAACCAACCGGCAACCTGATGTTTCTGAGAGAGAGAGGAGGAGCATCTCTGGAGG 3242
QY 1321 GAAAGTGATATATTTCCCGCTTTTAGATGGGTTTAAATAACTGATAATTTGATTTTATGATCA 1380

Db 3243 GAGGGCGATGATGTTCCCTTCTTGGACGGCTGATCAAGCCCTTATCGACTTCGACCCC 3302

Qy 1381 ATCCCAACACAGAAACAAATCTTTTGATGATCTCCAGCTTTTAATAATCTTCAACTAATAG 1440

Db 3303 GTGCCCCAACCAAGACCATCTTCGACGAGCAGCAGCGCGCGCGAGGAG 3362

Qy 1441 GAACAAACACTCCCCCGAATATCAGTTTAACTTTCTTTATTTCTGATGATGAGTTTGGAGT 1500

Db 3363 GACACAGCGCCAGCGCCCAACATCAGCTGACCTGAGCTACTTCCCAACATCAACGAG 3422

Qy 1501 GATACTGCCTACTCTCGGGAAACAGAGATGTTGTGATGACAGTTGAGGATTTGGAGT 1560

Db 3423 AACACCGCTTACAGCGCGAGAACAGAACGATCGACCGCGCTGAGGATCTGGAGC 3482

Qy 1561 GTGCAGGAGACGATTTGGCGGAGGCTTAGCTGGATACCATTTTTTGGSCCTGGAATC 1620

Db 3483 GTGCAGGAGACGACTGTGGCGCGGCTGAGCTGGATTCCTTTCTGGSCCGGCATC 3542

Qy 1621 GAAGGACTCTACTTCTGCGGTTTAAATCAAAATCAGAACTTAACTTTGTTAGTTGAGG 1680

Db 3543 GAGGGCTGTACACCGCGGCTTGATCAAGAACCGAACCAACCTGTGTGCGAGCTGAGG 3602

Qy 1681 CGCTTAGCTTAATCAAACTGCTAAATCCTTGGAGCTCTTGTTAAGGGTCACAAACCGAGGAA 1740

Db 3603 AGCTGCGCCAAACAGACCGCAGAGCTGGAGCTGCTGAGGGTGACCAACCGAGGAG 3662

Qy 1741 AGGACATTTTCTTAATCAATAGCGATGCAATTTGACTTTTTTGTCTTACGAGGTGGGCGGA 1800

Db 3663 AGGACCTTCAGCCTGATCAACAGGCAAGCATCGACTTCTGCTGACAGGTGGGCGGC 3722

Qy 1801 ACATGCAAGGTGCTAGGACCTGATTTGCTATGAGGATAGAGATCTATCTAAATATTC 1860

Db 3723 ACCTGCAAGGTGCTGGGCGGCGACTGCTGATCGGATCGAGGACCTGAGCAGGAACATC 3782

Qy 1861 TCAGAAACAAATCAGCAAAATCAGAAAGGATGAAACAAAAGGAGGAAACTGGCTGGGGTCTA 1920

Db 3783 AGCGAGCATCGACCGATCAAGAGGACGAGCAGAGGAGGCGCACCGGCTGGGCGCTG 3842

Qy 1921 GGTGGCAATGTGGACATCTGACTG 1946

Db 3843 GCGCGCAAGTGTGGACCGAGCTG 3868

RESULT 13

US-10-066-506A-1

Sequence 1, Application US/10066506A

Publication No. US20030108560A1

GENERAL INFORMATION:

APPLICANT: Grogan, Case C.

APPLICANT: Hevey, Michael C.

APPLICANT: Schmaljohn, Alan, L.

TITLE OF INVENTION: Chimeric Filovirus Glycoprotein

FILE REFERENCE: 003/243/SAP

CURRENT APPLICATION NUMBER: US/10/066,506A

CURRENT FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/267,522

PRIOR FILING DATE: 2001-01-31

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Apple Macintosh Microsoft Word 6.0

SEQ ID NO 1

LENGTH: 2252

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 624; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

US-10-066-506A-1

Query Match 26.7%; Score 546.6; DB 15; Length 2252;

Best Local Similarity 82.9%; Pred. No. 3.4e-151;

Matches 624; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

US-10-066-506A-1

Qy	1294	AGAAGAAACGATCGATTTTCTGGAAGAAGGTGATATATCCGGTTTTTAGATGGGTTA	1355
Db	1492	AGAACTCGACGATCGATTCCTCTGGAGGGAGGCGCATGTTCCCTTTTCTCGGATGGGTTA	1551
Qy	1354	ATAAATACGTGAAATTCGATTTTGGATCCAAATCCCAAAACACAGAAACAATCTTTGATGAATCT	1413
Db	1552	ATAAATGCTCCAATGATTTTGGACCCAGTTTCCAAATACAAAACAATCTTTGATGAATCC	1611
Qy	1414	CCGAGCTTTTAATACTTTCAACTAATGAGGAACAAACACATCCCCCGAATATCAGTTTAACT	1473
Db	1612	TCTAGTCTGGTCCCTCGGCTGAGGAAGATCAACATGCCTCCCCCAATATTAGTTAACT	1671
Qy	1474	TTCTCTATTATTTCTGATATAAAATGAGAGATATGCGCTACTCTGGGGAAAACGAGAAATGAT	1533
Db	1672	TTATCTATTATTTCTTAATAATAATGAGAACACATGCGCTACTCTGGAGAAAATGAGAAATGAT	1731
Qy	1534	TGATGACGAGTTGAGGATTTTCGAGTGTGTCAGAGGACGATTTTGGCGGCAGGCGTTAGC	1593
Db	1732	TGATGACGAGTTTAAAGATTTTGGAGCGTTTCAGAGATGACCTGGCGGCAGGCGTCACT	1791
Qy	1594	TGGATACCATTTTTTGGCCCTCGGAATCGAAGGACTCTATATCTGCCGGTTTTTAATCAAAAAAT	1653
Db	1792	TGGATACCGTTTTTGGCCCTCGGAATGGAAGGACTTTACACTGCTGTTTTTAATAAAAAAT	1851
Qy	1654	CAGAACAAATTTAGTTGTGTAGGTTGAGCGCTTAGCTAGTAATCAAACTGCTAAATCCTTTGGAG	1713
Db	1852	CAAAACAAATTTGGTCTGCGAGTTTGGCGGCTTAGCCCAATCAAACTGCAAAATCCTTTGGAA	1911
Qy	1714	CTCTGTTAAGGCTCACAACCGAGGAAGAGACATTTTCCCTTAATCAATAGGCATGCAAT	1773
Db	1912	CTCTTATTGAGAGTCAAACTGAGGAAGAAACATTTCTCTTAATCAATAGACATGCTATT	1971
Qy	1774	GACTTTTGTCTTACGAGGTGGGCGGGAACAATGCAAGGTGCTAGGACCTGATTTCTTGCGATA	1833
Db	1972	GACTTTCTACTCACAGATGGGAGGAGNACATGCAAGGTGCTTGGAATGATTTGATC	2031
Qy	1834	GGAAATGAGAGATCTATCTAAAAATATCTCAGAAACAAATCGACAAAATCGAAAAGGATGAA	1893
Db	2032	GGGATAGAGACTTGTCCAAAATAATTTACAGAGCAAAATTGACCAAAATTTAAAAGGACGAA	2091
Qy	1894	CAAAAGGAGAAACTGGCTGGGCTTAGGTGGGCAAAATGGTGGACATCTGACTGGGGTGTT	1953
Db	2092	CAAAAAGAGGGGACTGGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGTT	2151
Qy	1954	CTCACCAAATTTGGGCATCTGCTACTATTATCTATAGCTGTTCTGATTTGCTCTGCTCTCT	2013
Db	2152	CTTACTAACTTTGGCATTTTGCTACTATTATCTATAGCTGTTCTGATTTGCTCTCTCTCT	2211
Qy	2014	ATCTGTCGTATCTTCACTAAATACATTTGGATGA	2046
Db	2212	ATTGTGCTATCTTTACTAAATATATCGGATAA	2244
RESULT 14			
US-10-491-121-15			
; Sequence 15, Application US/10491121			
; Publication No. US20040259825A1			
; GENERAL INFORMATION:			
; APPLICANT: NABEL, GARY			
; APPLICANT: YANG, ZHI-YONG			
; APPLICANT: SULLIVAN, NANCY			
; APPLICANT: SANCHEZ, ANTHONY			
; TITLE OF INVENTION: Development of a preventive Vaccine for			
; TITLE OF INVENTION: Filovirus Infection in Primates			
; FILE REFERENCE: NIH221.001NP			
; CURRENT APPLICATION NUMBER: US/10/491,121			
; CURRENT FILING DATE: 2004-03-26			
; PRIOR APPLICATION NUMBER: PCT/US02/30251			
; PRIOR FILING DATE: 2002-09-24			
; PRIOR APPLICATION NUMBER: US 60/326476			
; PRIOR FILING DATE: 2001-10-01			
; NUMBER OF SEQ IDS NOS: 52			
; SOFTWARE: FastSeq for Windows Version 4.0			

```
; SEQ ID NO 15
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(S) delta TM
US-10-491-121-15

Query Match
Best Local Similarity 7.1%; Score 145; DB 20; Length 6940;
Matches 226; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1530 TGAATGTGATGAGAGTTGGAGATTGGAGTGTGAGAGGAGGATTTGGCGGCGAGGCT 1589
Db 3483 TAAATGCAATCCAACTTACACTACTGGACTGCACAAGAACAAACATAATGTCTGGGAT 3542
QY 1590 TAGCTGGATACCATTTTGGCCCTGGAATCGAAGGACTCTATCTGCGGTTTAAATCAA 1649
Db 3543 TGCCTGGATCCCGTACTTTGGACCGGGTGCAGAAGGCATATACCTGAAGGCCCTTATGCA 3602
QY 1650 AAATCAGAACAAATTTAGTTTGTAGGTTGAGGCGCTTAGCTAATCAAACTGCTAAATCCTT 1709
Db 3603 CAACCAAAATGCTTAGTCTGGACTCAGACAACTTGCMAATGAACAACTCAAGCTCT 3662
QY 1710 GGAGCTCTTTTAAAGGTCACAAACGAGAAAGGACATTTTCCCTTAATCAATAGGCATGC 1769
Db 3663 GCAGCTTTTCTTAAAGGCGCACGACGAGCTGGGACATATACCATCTCAATAGGAAGGC 3722
QY 1770 AATTGACTTTTGTCTACGAGTGGGCGGACATCGAAGGTCAGGACCTGATGTTG 1829
Db 3723 CATAGATTTCTTCTGCGAGATGGGCGGACATGTAGGATCCTGGGACCAAGATTGTTG 3782
QY 1830 CATAGGAATAGAGATCTATCTTAAATAATCTCAGAACAAATCGACAAATCAGAAAGGA 1889
Db 3783 CATTGAGCCACATGATTGGACCAAAACATCACTGATAAATCAACCAATCATCCATGA 3842
QY 1890 T 1890
Db 3843 T 3843

RESULT 15
US-10-860-878-2
; Sequence 2, Application US/10860878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
; FILE REFERENCE: 2115s-001474POA
; CURRENT APPLICATION NUMBER: US/10/860, 878
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/09/600, 766
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: USSN 60/072033
; PRIOR FILING DATE: 1998-01-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 7073
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; OTHER INFORMATION: containing DNA for GP of Ebola Virus, Sudan strain
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (886)..(1129)
; OTHER INFORMATION: CMV IE 5' UT
; FEATURE:
; NAME/KEY: Intron
```

```
; LOCATION: (1130)..(1840)
; OTHER INFORMATION: CMV IE INT
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: (248)..(885)
; OTHER INFORMATION: CMV enhancer
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1870)..(4089)
; OTHER INFORMATION: GP(S)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4090)..(4642)
; OTHER INFORMATION: TDGH
; FEATURE:
; NAME/KEY: gene
; LOCATION: (6138)..(6760)
; OTHER INFORMATION: Kan r
; US-10-860-878-2
```

```
Query Match
Best Local Similarity 7.1%; Score 145; DB 22; Length 7073;
Matches 226; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1530 TGAATGTGATGAGAGTTGGAGATTGGAGTGTGAGAGGAGGATTTGGCGGCGAGGCT 1589
Db 3503 TAAATGCAATCCAACTTACACTACTGGACTGCACAAGAACAAACATAATGTCTGGGAT 3562
QY 1590 TAGCTGGATACCATTTTGGCCCTGGAATCGAAGGACTCTATCTGCGGTTTAAATCAA 1649
Db 3563 TGCCTGGATCCCGTACTTTGGACCGGGTGCAGAAGGCATATACCTGAAGGCCCTTATGCA 3622
QY 1650 AAATCAGAACAAATTTAGTTTGTAGGTTGAGGCGCTTAGCTAATCAAACTGCTAAATCCTT 1709
Db 3623 CAACCAAAATGCTTAGTCTGGACTCAGACAACTTGCMAATGAACAACTCAAGCTCT 3682
QY 1710 GGAGCTCTTTTAAAGGTCACAAACGAGAAAGGACATTTTCCCTTAATCAATAGGCATGC 1769
Db 3683 GCAGCTTTTCTTAAAGGCGCACGACGAGCTGGGACATATACCATCTCAATAGGAAGGC 3742
QY 1770 AATTGACTTTTGTCTACGAGTGGGCGGAAACATCGAAGGTCAGGACCTGATGTTG 1829
Db 3743 CATAGATTTCTTCTGCGACGATGGGCGGACATGTAGGATCCTGGGACCAAGATTGTTG 3802
QY 1830 CATAGGAATAGAGATCTATCTTAAATAATCTCAGAACAAATCGACAAATCAGAAAGGA 1889
Db 3803 CATTGAGCCACATGATTGGACCAAAACATCACTGATAAATCAACCAATCATCCATGA 3862
QY 1890 T 1890
Db 3863 T 3863
```

Search completed: August 1, 2005, 20:41:05
Job time : 1291.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 31, 2005, 13:57:58 ; Search time 8793.18 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatatattttct.....ttactaaatatatggataa 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2046	100.0	2046	6	AX717719	Sequence
2	1847.6	90.3	2046	6	AX717725	Sequence
3	1842.8	90.1	2046	14	AF005734	Marburg v
4	1514.8	74.0	11460	6	AR279677	Sequence
5	1514.8	74.0	19104	14	WVREPCYC	212132 Marburg vir
6	1513.2	74.0	19112	14	AY430366	AY430366 Lake Vict
7	1513.2	74.0	19113	14	AY430365	AY430365 Lake Vict
8	1485.6	72.6	2051	6	AX717721	Sequence
9	1474.8	72.1	2046	14	AF005735	Marburg v
10	1473.2	72.0	2948	14	MAVSPA	68493 Marburg Vir
11	1473.2	72.0	19112	14	MVIRPR	239337 Marburg vir
12	1443.4	70.5	19151	14	AY358025	Marburg v
13	1441.2	70.4	2046	14	AF005733	Marburg v
14	1319.6	64.5	2046	6	AX717717	Sequence
15	780	38.1	1841	6	AX717715	Sequence
16	746.6	36.5	2252	6	AX717713	Sequence
17	149	7.3	18890	14	AB050936	Reston Eb
18	148.2	7.2	2408	14	EVU28006	Cote d'Ivoi
19	145.8	7.1	2033	14	EVU23416	Reston Ebol

20	145.8	7.1	2033	14	EVU23417	U23417 Reston Ebol
21	145.8	7.1	2233	14	AF034645	Ebola vir
22	145.8	7.1	2360	14	EVU23152	U23152 Reston Ebol
23	145.8	7.1	18891	14	AF522874	AF522874 Reston Eb
24	139.2	6.8	2261	14	AY344234	AY344234 Sudan ebo
25	139.2	6.8	2363	14	AY316199	AY316199 Sudan Ebo
26	139.2	6.8	18875	14	AY729654	AY729654 Sudan ebo
27	137.2	6.7	2408	14	EVU81161	U81161 Zaire Ebola
28	135.6	6.6	18959	14	AY142960	AY142960 Zaire Ebo
29	135.6	6.6	18960	14	AF499101	AF499101 Zaire Ebo
30	134.4	6.6	2362	14	EVU23069	U23069 Sudan Ebola
31	134.4	6.6	2362	14	EVU28134	U28134 Sudan Ebola
32	134	6.5	2039	6	AX717723	Sequence
33	134	6.5	2298	6	AR404895	AR404895 Sequence
34	134	6.5	2298	6	AX092102	AX092102 Sequence
35	134	6.5	2406	14	EVU31033	U31033 Zaire Ebola
36	134	6.5	2408	14	EVU23187	U23187 Zaire Ebola
37	134	6.5	11742	14	EBORNA	L11365 Zaire Ebola
38	134	6.5	18959	14	AF086833	AF086833 Zaire Ebo
39	134	6.5	18959	14	AF272001	AF272001 Zaire Ebo
40	130.8	6.4	2408	14	EVU28077	U28077 Zaire Ebola
41	130.8	6.4	18961	14	AY354458	AY354458 Zaire ebo
42	127.6	6.2	2173	14	AY058898	AY058898 Zaire Ebo
43	126	6.2	2173	14	AY526098	AY526098 Zaire ebo
44	126	6.2	2173	14	AY526100	AY526100 Zaire ebo
45	126	6.2	2173	14	AY526101	AY526101 Zaire ebo

ALIGNMENTS

RESULT 1
AX717719
LOCUS AX717719 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 7 from Patent WO02079239.
ACCESSION AX717719
VERSION AX717719.1 GI:29890728

KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 7 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)

FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Raven Glycoprotein 1 and Marburg virus strain Musoke
Glycoprotein 2"

ORIGIN

Query Match 100.0%; Score 2046; DB 6; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60
Qy 61 GTTTTAGAAATTCGTAGTAACAGCCACCTCAAGATGATGATTCAGTGTGTCGCGGAACC 120
Db 61 GTTTTAGAAATTCGTAGTAACAGCCACCTCAAGATGATGATTCAGTGTGTCGCGGAACC 120
Qy 121 CTCCAAAGACAGAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCCAAAGACAGAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCTTTGGAGAGCATCTAAAGATGGGCTTTCAGGACAGGTTCTCTCCCCAGAAC 240

Db 181 ||||| GATTCCTCCCTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTTCTCCCTCCCAAGAAC 240
QY 241 GTTGGAGTATACGGAAGGAGAGCAAGCCAAACATGTTACAATATAAAGTGTAACGACCCCT 300
Db 241 GTTGGAGTATACGGAAGGAGAGCAAGCCAAACATGTTACAATATAAAGTGTAACGACCCCT 300
QY 301 TCTGAAATCCTTGCTGCTGGATCCTCCAGTAATATCCGCGATTACCCCTAAATGTAAA 360
Db 301 TCTGAAATCCTTGCTGCTGGATCCTCCAGTAATATCCGCGATTACCCCTAAATGTAAA 360
QY 361 ACTGTTATCATATTTCAAGGTCAAAACCTCATGACAGGGGATGGCCCTCGATTGTGG 420
Db 361 ACTGTTATCATATTTCAAGGTCAAAACCTCATGACAGGGGATGGCCCTCGATTGTGG 420
QY 421 GGGGCAATTTTCTGTATGATCGCGTGCCTCTCAACAATGTACCGAGGCAAGGTCCTC 480
Db 421 GGGGCAATTTTCTGTATGATCGCGTGCCTCTCAACAATGTACCGAGGCAAGGTCCTC 480
QY 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGACAGTTTCACAGAATGATTTTTCT 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGACAGTTTCACAGAATGATTTTTCT 540
QY 541 AGGCAAGGACAAAGGTTATCGTCACATGAACTTTGACCTCCACCAATAAATATTTGGACAAGC 600
Db 541 AGGCAAGGACAAAGGTTATCGTCACATGAACTTTGACCTCCACCAATAAATATTTGGACAAGC 600
QY 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTGGCATCCTCCAGAAATACAACTCC 660
Db 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTGGCATCCTCCAGAAATACAACTCC 660
QY 661 ACAACAATCAAAATGATGCTCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
Db 661 ACAACAATCAAAATGATGCTCTCTTAAACCTCCATCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
QY 721 AGCATTTCACTCTCAAAATCTCAAAATTAATCTGCTAAATCTGGAACATATGAACCAAGT 780
Db 721 AGCATTTCACTCTCAAAATCTCAAAATTAATCTGCTAAATCTGGAACATATGAACCAAGT 780
QY 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGACAGGGGCCACACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGACAGGGGCCACACA 840
QY 841 ACTCTTAATGTAGTCACTGAAAGAAACAATCGTCAACAATATTTGTCACCTCTTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAAAGAAACAATCGTCAACAATATTTGTCACCTCTTCACTA 900
QY 901 CATCAAGACCTCAACATGAGCAAAACAGTACGAATCCTTCCCGACATGCTGTAAC 960
Db 901 CATCAAGACCTCAACATGAGCAAAACAGTACGAATCCTTCCCGACATGCTGTAAC 960
QY 961 GAGCAATGGAACCGACCCCAACAACAACAGCAACGCTCTCAACAATACTAATACA 1020
Db 961 GAGCAATGGAACCGACCCCAACAACAACAGCAACGCTCTCAACAATACTAATACA 1020
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTTCCCTCCCAACCGCAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTTCCCTCCCAACCGCAAC 1080
QY 1081 ATCAACAATATGATACAAAGTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
Db 1081 ATCAACAATATGATACAAAGTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
QY 1141 AACCAACTCTAGATCCCAAGAAATCCCAACAGGACAGACCAACAGCAACACC 1200
Db 1141 AACCAACTCTAGATCCCAAGAAATCCCAACAGGACAGACCAACAGCAACACC 1200
QY 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACCCCAAAATTTCTTCGGAT 1260
Db 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACCCCAAAATTTCTTCGGAT 1260
QY 1261 TCTAGTCGCAACCGCGCTCTTATATATCTTTAGAAAGAACGATCGATCCTCTGGAGG 1320
|||||

Db 1261 TCTAGTCGCAACCGCGCTCTTATATATCTTTAGAAAGAAACGATCGATCCTCTGGAGG 1320
QY 1321 GAAGCGCACATGTTCCCTTTCTCGATGGGTAAATAATGCTCCAAATGATTTTGGACCA 1380
Db 1321 GAAGCGCACATGTTCCCTTTCTCGATGGGTAAATAATGCTCCAAATGATTTTGGACCA 1380
QY 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTGCTGCTCGGCTGAGGAA 1440
Db 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTGCTGCTCGGCTGAGGAA 1440
QY 1441 GATCAACATGCTCCCTCCCAATATTAGTTTAACCTTTATCTTATTTTCTTAATAATAAG 1500
Db 1441 GATCAACATGCTCCCTCCCAATATTAGTTTAACCTTTATCTTATTTTCTTAATAATAAG 1500
QY 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAAGAAATTTGGAGC 1560
Db 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAAGAAATTTGGAGC 1560
QY 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCCGTTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCCGTTTTTTTGGCCCTGGAAAT 1620
QY 1621 GAAGGACTTTTACCTGCTGTTTTTAATAATAATAATAATAATAATAATAATAATAATA 1680
Db 1621 GAAGGACTTTTACCTGCTGTTTTTAATAATAATAATAATAATAATAATAATAATAATA 1680
QY 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTTGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTTGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
QY 1741 AGAATCTCTCTTAAATCAATAGACATGCTATTGACTTTTACTCAAGAATGGGAGGA 1800
Db 1741 AGAATCTCTCTTAAATCAATAGACATGCTATTGACTTTTACTCAAGAATGGGAGGA 1800
QY 1801 ACATGCAAGTCTTGACCTGATTGTCATCGGATAGAGACTTGTCCAAAATATTT 1860
Db 1801 ACATGCAAGTCTTGACCTGATTGTCATCGGATAGAGACTTGTCCAAAATATTT 1860
QY 1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAGAGGGGACTGGTTGGGCTG 1920
Db 1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAGAGGGGACTGGTTGGGCTG 1920
QY 1921 GGTGTAATAGTGGACATCGACTGGGGTGTCTTAACTAACTGGGCAATTTTGGCTACTA 1980
Db 1921 GGTGTAATAGTGGACATCGACTGGGGTGTCTTAACTAACTGGGCAATTTTGGCTACTA 1980
QY 1981 TTATCCATAGTCTTGTGATGCTCTATCTGATTTGCTGATTTGCTGATTTTACTAATATATC 2040
Db 1981 TTATCCATAGTCTTGTGATGCTCTATCTGATTTGCTGATTTTACTAATATATATC 2040
QY 2041 GGATAA 2046
Db 2041 GGATAA 2046
RESULT 2
AX717725
LOCUS AX717725 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 13 from Patent WO02079239.
ACCESSION AX717725
VERSION AX717725.1 GI:29890731
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 13 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source Location/Qualifiers
1. .2046

ORIGIN

Query Match 90.3%; Score 1847.6; DB 6; Length 2046;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

1 ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAAACTCTCCCT 60

61 GTTTTAGAAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
Db 61 GTTTTAGAAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120

121 CTCMAAGACAGAAGATGTTTCATCTGTATGGGATTTTACATGATGGGCAAAAAGTTGCT 180
Db 121 CTCMAAGACAGAAGATGTTTCATCTGTATGGGATTTTACATGATGGGCAAAAAGTTGCT 180

181 GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTTCTCCCAAGAAC 240
Db 181 GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTTCTCCCAAGAAC 240

241 GTTGAGTATACGGAAGGAGAGAGCAAAACATGTTTACAATATAAGTGTAAACAGACCT 300
Db 241 GTTGAGTATACGGAAGGAGAGAGCAAAACATGTTTACAATATAAGTGTAAACAGACCT 300

301 TCTGMAAATCTTGCTGCTGATCTCCAGATTAATTCGCGATTTACCTTAAATGTAAA 360
Db 301 TCTGMAAATCTTGCTGCTGATCTCCAGATTAATTCGCGATTTACCTTAAATGTAAA 360

361 ACTGTTTCATATATTCAGGTTCAAAACCTCATGCACAGGGGATGTCCTCCATTTGG 420
Db 361 ACTGTTTCATATATTCAGGTTCAAAACCTCATGCACAGGGGATGTCCTCCATTTGG 420

421 GGGGCAATTTTCTGTATGATCGCGTTCCTCTACAACATGTACCGAGGCAAGTCTTC 480
Db 421 GGGGCAATTTTCTGTATGATCGCGTTCCTCTACAACATGTACCGAGGCAAGTCTTC 480

481 ACTGAAGGAAATATAGCAGCTATGATTTTAAAGACAGTTTCAAGAAATGATTTTCT 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTTAAAGACAGTTTCAAGAAATGATTTTCT 540

541 AGGCAAGGACAGGTTATCGTCACATGAATCTTGACCTCCACCAATATATGGACAGC 600
Db 541 AGGCAAGGACAGGTTATCGTCACATGAATCTTGACCTCCACCAATATATGGACAGC 600

601 AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCTCAAGAAATACAACTCC 660
Db 601 AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCTCAAGAAATACAACTCC 660

661 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAATCCG 720
Db 661 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAATCCG 720

721 AGCATTCATCTACAAATTAATTAATCTGCTAAATCTGGAATCTATGAAACCAAGT 780
Db 721 AGCATTCATCTACAAATTAATTAATCTGCTAAATCTGGAATCTATGAAACCAAGT 780

781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCACACA 840

841 ACTCTTAATGTAGTCACTGAACAGAAACAACTCGTCAACAATATTTGTCACCTCTTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAACAGAAACAACTCGTCAACAATATTTGTCACCTCTTCACTA 900

901 CATCCAGCACCTCAACATGAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAAC 960

901 CATCCAGCACCTCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
961 GAGCACAATGGAAACCGACCCCAACAAACAGCAGCAACGCTCTCAACAATACTAATAACA 1020
961 GAGCACAATGGAAACCGACCCCAACAAACAGCAGCAACGCTCTCAACAATACTAATAACA 1020

1021 ACTCCGACCTTAACACTCTCAAGTCAACCTCAGTACTCTTCCCTCCCAACCGGCAAC 1080
Db 1021 ACTCCGACCTTAACACTCTCAAGTCAACCTCAGTACTCTTCCCTCCCAACCGGCAAC 1080

1081 ATCAACCAATATGATACAAACAGTGAATAGCAGAAACGCAAAACAAACAAATGCTCAGTTG 1140
Db 1081 ATCAACCAATATGATACAAACAGTGAATAGCAGAAACGCAAAACAAACAAATGCTCAGTTG 1140

1141 AACCAACTCTTAGATCCAAACAGAAATATCCCAACAGGACAGACACCAACAGCAACACC 1200
Db 1141 AACCAACTCTTAGATCCAAACAGAAATATCCCAACAGGACAGACACCAACAGCAACACC 1200

1201 AACATCATGACGACATCAGATATAACAGCAAAACACCCCAACAAATTTCTTCCGGAT 1260
Db 1201 AACATCATGACGACATCAGATATAACAGCAAAACACCCCAACAAATTTCTTCCGGAT 1260

1261 TCTAGTCCGACAAACCGCTCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 1261 TCTAGTCCGACAAACCGCTCTCTATATCTTTAGAAAGAAACGATCGATTTCTGGAAA 1320

1321 GAAGGCGACATGTTTCCCTTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGACCCA 1380
Db 1321 GAAGGCGATATATTTCCCTTTTCTAGATGGGTAAATAATACTGAAATTTGATTTGATCCA 1380

1381 GTTCCAAATACAAAACAACTTTTGATGAATCTCTAGTCTGGTCTCGGCTCGGAGAA 1440
Db 1381 ATCCCAACACAGAAACAACTTTTGATGAATCTCTCGAGCTTAACTACTTCAACTAATGAG 1440

1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATAATAACAG 1500
Db 1441 GAACACACATCTCCCGAATATCAGTTTAACTTTTCTTATTTCTTGATTAATAANTGGA 1500

1501 AACACTGCTACTCTCGGAGAAATGAGAAATGATTTGATGAGAGTTTAAAGAAATTTGGAGC 1560
Db 1501 GATACTGCTACTCTCGGAGAAACGAGAAATGATTTGATGAGAGTTTGGAGTTTGGAGT 1560

1561 GTTCAGAGAGATGATCTGGCCGAGGGCTCAGTTGGATACCGTTTTTGGCCCTCGGAAT 1620
Db 1561 GTGCAGAGAGAGATTTGGCCGAGGGCTTAGCTGGATACCAATTTTGGCCCTCGGAATC 1620

1621 GAAGGACTTTACACTGCTGTTTTTAAATAAATAACAAATTTGGTCTCGAGGTTGAGG 1680
Db 1621 GAAGGACTTTATATCTGCGGTTTTAATAAATAACAAATTTAGTTTGTAGTTGAGGTTGAGG 1680

1681 CGTCTAGCCAATCAAACTGCCAAATCTTTGGAACTCTTTATTTGAGAGTCAACCTGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTAAGGTCACAACTGAGGAA 1740

1741 AGAACAATTTCTCTTAATCAATAGACATGCTATGATCTTCTACTCAAGATGGGGAGA 1800
Db 1741 AGGACATTTTCTCTTAATCAATAGGATGCAATTTGACTTTTTTGTACGAGGTTGGGGCGGA 1800

1801 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAACTTCTTCCAAAATAT 1860
Db 1801 ACATGCAAGGTTGCTAGGACCTGATTTGTCATAGAAATAGAAATCTATCTAAAATATC 1860

1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGCAACAAAGAGGGGACTTGGTTGGGGCTG 1920
Db 1861 TCAGAACAAATCGACAAATTCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGGCTA 1920

1921 GGTGGTAAATGGTGGACATCCGACTGGGGTCTTCTACTAATCTTGGGCAATTTGCTACTA 1980
Db 1921 GGTGGCAAAATGGTGGACATCTGACTGGGGTCTTCTCAACAAATTTGGGCACTCTGCTACTA 1980

1981 TTATCCATAGCTCTTGAATGCTCTATCTCGTATTTGTCGATCTTCTTACTTAAATATATC 2040

Db 1981 TTATCTATAGCTGTTCTGATTGCTCTGCTCTGCTATCTGCTGCTATCTTCACTAAATACATT 2040
QY 2041 GGATAA 2046
Db 2041 GGATGA 2046
RESULT 3
AF005734
LOCUS
DEFINITION Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein precursor (GP) gene, complete cds.
ACCESSION AF005734
VERSION AF005734.1 GI:2459877
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PUBMED 9448698
REFERENCE 2 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
source
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Kenya/Kitum Cave/1987/Ravn"
/db_xref="taxon:11269"
1..2046
/gene="GP"
1..2046
/gene="GP"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40459.1"
/db_xref="GI:2459878"
/translation="MKTIYFLISLILIQIKTLPVLEIASNSQPDVDSVCSGLQKT EDVHLMGFTLSGQKVDSPLEASKRWAFRTGVPVKNVEYTEGEAKTCYNI SVTDPG KSLIDPPSNIRDYPKCTVHHIOQNPHAQGIALHMGAFFLYDRVASTTMYRGKVF TEGNTAAMI VNKTVHRMI FSRQGGYRHMNLTSTNKYWTSSNETORNDTGCFGILOEY NSTNNQTCPSLSKPSELPVTPSIHSTNTQINTAKSGTMNPSSDDEDLMSGSGEQ GPHTLVNVTVQKQSSTILSPSLPSSQHEQNSTNPSRHAHVHEHNGTDFTTPQAIL LNNNTTYNTLKNLSTPPPTNNDTQRELAESQTNALQNTLTDPTENPT GQDNTSTNIIMTSDITSKPTNSSPDSPTTPPIYIFKRKRSIFWKEGDIFPFLDG LINTIDEPPIPTNETIFDESFSNTSNEOHTPPNLSLTFSPDKNGDTPAYSGEN ENDCDAELR IKSVDSDLAAGLSMIPFPGPIEGLYTAGL IKNQNNLVCRLLRANOT AKSLELLRVTEERTFSLNRHAI DFLITWGTCKVLGPDCCIGIEDLSKNISEQI DKIRDEQKEETGWTGGKRWTS'DMGVLTNLGILLLSIAVLIALSCICRIFTKIY"
1..81
sig_peptide
/gene="GP"
mat_peptide
82..2043
/gene="GP"
/product="glycoprotein"
misc_feature
1930..2019
/gene="GP"
/note="encodes transmembrane anchor peptide"
ORIGIN
Query Match 90.1%; Score 1842.8; DB 14; Length 2046;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1919; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAATCCAAAGTATATAAAACTCTCCCT 60
Db 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAATCCAAAGTATATAAAACTCTCCCT 60
QY 61 GTTTTAGAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTGAGTGTGCTCCGGAACC 120
Db 61 GTTTTAGAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTGAGTGTGCTCCGGAACC 120
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTTGGAAAGCATCTTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTTGGAAAGCATCTTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 240
QY 241 GTTGAGTATACGGAAGGAGAAGAGCCAAACATGTTACAATATAAGTGTAAACAGACCT 300
Db 241 GTTGAGTATACGGAAGGAGAAGAGCCAAACATGTTACAATATAAGTGTAAACAGACCT 300
QY 301 TCTGAAAAATCCTTGTCTGCTGGATCCTCCCAAGTAAATATCCCGGATTTACCTTAAATGTAAA 360
Db 301 TCTGAAAAATCCTTGTCTGCTGGATCCTCCCAAGTAAATATCCCGGATTTACCTTAAATGTAAA 360
QY 361 ACTGTTTCATCATATTTCAAGGTCAAACCCCTCATGCACAGGGGGATTGCCCTCCATTGTGG 420
Db 361 ACTGTTTCATCATATTTCAAGGTCAAACCCCTCATGCACAGGGGGATTGCCCTCCATTGTGG 420
QY 421 GGGGCATTTTCTGTATGATCGGGTTGCCCTCTCAACAATGTACCGAGGCAAGTCTTC 480
Db 421 GGGGCATTTTCTGTATGATCGGGTTGCCCTCTCAACAATGTACCGAGGCAAGTCTTC 480
QY 481 ACTCAAGGAATATAGCAGCTATGATTGTTAATAAGCAGTTTACAGAAATGATTTTTCT 540
Db 481 ACTCAAGGAATATAGCAGCTATGATTGTTAATAAGCAGTTTACAGAAATGATTTTTCT 540
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCCAACAATAATTTGGCAAGC 600
Db 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCCAACAATAATTTGGCAAGC 600
QY 601 AGCAATGAACGAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAATACAACTCC 660
Db 601 AGCAATGAACGAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAATACAACTCC 660
QY 661 ACAACCAATCAACATCGCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAAGTCCG 720
Db 661 ACAACCAATCAACATCGCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAAGTCCG 720
QY 721 AGCATTCACTCTCAAAATACTCAAATTAATACTGCTAAATCTGGAACCTATGAACCCAAGT 780
Db 721 AGCATTCACTCTCAAAATACTCAAATTAATACTGCTAAATCTGGAACCTATGAACCCAAGT 780
QY 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGATCTGAGAAACAGGGGCCCCACACA 840
Db 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAAACAGGGGCCCCACACA 840
QY 841 ACTCTTAATGTAGTCACTGAAACAAACAAATCGTCAACAATATTGTCCACCTCTTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAAACAAATCGTCAACAATATTGTCCACCTCTTCACTA 900
QY 901 CATCCAAGCACCTCACAACATGAGCAAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
Db 901 CATCCAAGCACCTCACAACATGAGCAAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
QY 961 GAGCACAATGGAAACCGACCCCAACAACAACAGCAGCAAGCTCTCTCAAACTACTAATACA 1020
Db 961 GAGCACAATGGAAACCGACCCCAACAACAACAGCAGCAAGCTCTCTCAAACTACTAATACA 1020
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCGCAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCGCAAC 1080

```
Qy 1081 ATCACTAATATGATACACACGCTGAACCTAGCAGAAAGCGAACCAAAACCAATGCTCAGTTG 1140
Db 1081 ATCACTAATATGATACACACGCTGAACCTAGCAGAAAGCGAACCAAAACCAATGCTCAGTTG 1140
Qy 1141 AACCAACTCTAGATCCAAACAGAAATCCCAACAGCAGCAAGACCAACAGCAGCAAC 1200
Db 1141 AACCAACTCTAGATCCAAACAGAAATCCCAACAGCAGCAAGACCAACAGCAGCAAC 1200
Qy 1201 AACATCATGACGACATCAGATATAAACAAGCAAAACCAACCAAAATCTTCTCCGAT 1260
Db 1201 AACATCATGACGACATCAGATATAAACAAGCAAAACCAACCAAAATCTTCTCCGAT 1260
Qy 1261 TCTAGTCGACAAACCCGCTCTCTATATACCTTATAGAAAGAAACGATCGATCTCTGGAG 1320
Db 1261 TCTAGTCGACAAACCCGCTCTCTATATACCTTATAGAAAGAAACGATCGATCTCTGGAG 1320
Qy 1321 GAAGCGCATGTTCCCTTTCTCGATGGGTTAATAATGCTCCAATTTGATTTGACCCA 1380
Db 1321 GAAGGTGATATATCCCGTTTATAGATGGTTAATAATGCTGAAATTTGATTTGATCCA 1380
Qy 1381 GTTCCAATACAAAACAACTCTTGATGAATCCTCTAGTCTGGTGCTCGGCTGAGGAA 1440
Db 1381 ATCCCAACACAGAAACAACTCTTGATGAATCCTCCAGCTTTAATACTTCAACTAATGAG 1440
Qy 1441 GATCAACATGCTCTCCCAATATAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Db 1441 GAACAACACATCTCCCGCAATATCAGTTTAACTTTCTTATTTCTGATATAAATGGA 1500
Qy 1501 AACCTGCTACTCTGAGAAATAGAGAAATGATGATGATGAGAGTTAAGAAATTTGGAGC 1560
Db 1501 GATCTGCTACTCTGAGGAAACGAGAAATGATGATGATGAGAGTTGAGATTTGGAGT 1560
Qy 1561 GTTCAGGAGATGACTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTTCAGGAGGAGATTTGGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
Qy 1621 GAAGACTTTACATCTGCTGTTTTTAAATAAATCAAAACAAATTTGCTGAGTTGAGG 1680
Db 1621 GAAGACTCTTACTGCTGCGGTTTTTAAATAAATCAAAATCAGAAATTTAGTTTGGTTGAGG 1680
Qy 1681 CGCTAGCAATCAAACTGCCAAATCTCTTGGAACTCTTATTTAGAGAGTCAAACTGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCAAAACCGAGGAA 1740
Qy 1741 AGAACATTTCTTAATCAATAGACATGCTATGACTTTCTAATCAAGATGGGAGGA 1800
Db 1741 AGGACATTTCTTAATCAATAGGATGCAATTTGACTTTTCTTACGAGGTGGGCGGA 1800
Qy 1801 ACATGCAAGTCTTGGACCTGATTTGATCGGATAGAGACTTTGTCCAAAATATT 1860
Db 1801 ACATGCAAGGTCTAGGACCTGATTTGATGATAGGAATAGGAATCTATCTTAAAAATATC 1860
Qy 1861 TCAGAGCAATTTGACCAATTAAGAAAGCAACAAAGAGGAGGACTGTTGGGCTG 1920
Db 1861 TCAGAACAAATCGAACAAATCAGAAAGATGAACAAAGAGGAGAACTGGCTGGGCTGA 1920
Qy 1921 GGTGTAATGTTGGACATCCGACTGGGCTGTTCTTAACTTTGGGCAATTTTGTCTACTA 1980
Db 1921 GGTGCAATGTTGGACATCTGACTGGGCTGTTCTCACCATTTGGGCAATCTGCTACTA 1980
Qy 1981 TTATCCATAGCTGCTTGAATGCTCTATCTGATATTTGCTGATCTTTTACTAATAATATC 2040
Db 1981 TTATCTATAGCTGTTCTGATGCTCTGCTGATCTGCTGATCTGCTGATCTTCACTAAATATC 2040
Qy 2041 GGATGA 2046
Db 2041 GGATGA 2046
```

RESULT 4

AR279677

LOCUS

DEFINITION

Sequence 1 from patent US 6517842.

11460 bp

DNA

linear

PAT 10-APR-2003

ACCESSION AR279677

VERSION AR279677.1

KEYWORDS GI:29714600

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 11460)

AUTHORS Hevey,M.C., Negley,D.L., Pushko,P., Smith,J.F. and Schmaljohn,A.L.

TITLE Marburg virus vaccines

JOURNAL Patent: US 6517842-A 1 11-FEB-2003;

FEATURES Location/Qualifiers

1..11460

source /organism="unknown"

/mol_type="genomic DNA"

ORIGIN

```
Query Match 74.0%; Score 1514.8; DB 6; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 5940 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 5999
Qy 61 GTTTTAGAATTTCTAGTACAGCAACCTCAAGATGTAGATTAGTGTGCTCCGGAAC 120
Db 6000 ATTTTAGAGATAGTGTGTAATATCAACCCCAAAATGTGGATTCGGTATGTCTCGGAAC 6059
Qy 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTTACACTGAGTGGGCAAAAAGTTGCT 6119
Qy 181 GATTCCCTTTTGGAAAGCATCTAAACGATGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 240
Db 6120 GATTCCCTTTTGGAGGATCTCAAGCATGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 6179
Qy 241 GTTCAGTATACGGAAGGAGAGACCAACATGTTACATATAAGTGTACAGACCT 300
Db 6180 GTTCAGTATACGGAAGGAGAGACCAACATGTTACATATAAGTGTACAGACCT 6239
Qy 301 TCTGAAAATCTTGTGCTGGATCTCCCAAGTAAATATCCCGATTTACCTTAAATGTAAA 360
Db 6240 TCTGAAAATCTTGTGCTGGATCTCCCAAGTAAATATCCCGATTTACCTTAAATGTAAA 6299
Qy 361 ACTGTTTCATATATTCAGGTCAAAACCCCTCATGCAAGGGATTTGCCCTCATTTGTGG 420
Db 6300 ACTATCCATCATATTCAGGTCAAAACCCCTCATGCAAGGGATTCGCCCTCATTTATGG 6359
Qy 421 GGGCATTTTTTCTGTATGATCGCGTTGCTCTTACCAACATGTACCGAGGCAAGTCTTC 480
Db 6360 GGGCATTTTTTCTGTATGATCGCGATTTGCCCTTCCCAACATGTACCGAGGCAAAAGTCTTC 6419
Qy 481 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGATGATTTTTTCT 540
Db 6420 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGATGATTTTTTCT 6479
Qy 541 AGGCAAGCAAGGTTATTCGTCACTGAACTTGACCTCCCAACATTAATTTTGGACAAGC 600
Db 6480 CGGCAAGCAAGGTTATTCGTCACTGAACTTGACCTCCCAACATTAATTTTGGACAAGT 6539
Qy 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC 660
Db 6540 AGTAAACGAAACGCAAAACGAATGACATGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 6599
Qy 661 ACAACCAATCAACATGCCCTCCATCTTTAAACCTCCATCCCTGCCCAACAGTAACTCCG 720
Db 6600 ACAAGCAACCAACATGCTGCTCCGTCCAAAATACCTCCCAACCTCCCAACAGTAACTCCG 6659
Qy 721 AGCAATCACTCTACAAATTAATTAATCTGTAATCTGGAATCTGGAATCTGGAATCTGGA 780
Db 6660 GAGATCAAACTCAACAGCAACCACTGATGTCACCAACCACTCAATACCAAGCAACCAAGC 6719
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
```

Db 6720 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGACGAGAAACCCACACA 6779
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAAACAATCGTCAACAATATGTTCCACTCTTCACTA 900
Db 6780 ACTTCTGATGCGGTCAACGACGAGGGCTTTCATCAACATGCGACCCACTCCCTCACA 6839
Qy 901 CATCAAGCACCTCAACAATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAACT 960
Db 6840 CAACCAAGCAGCCACAGCAGGAGGAAACAACAACAACCAATTCGCCAAGATGCTGTGACT 6899
Qy 961 GAGCAATGGAACCGACCCCAACAACAACAACGACCAACGCTCTCTCAACAATACTAATACA 1020
Db 6900 GAACTAGACAAAATAACACAACTGCAACACGCTCCATGCCCTCCATCAACACTACCAACA 6959
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCGCAAC 1080
Db 6960 ATCTCTACTAACAACACCTCCAAACAACAATTCAGCACTCTCTGACACCATTAACAAC 7019
Qy 1081 ATCAACAATATGATACACAACGTAAGCTAGCAGAAAGCGAAACAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACAACACAGAGCACAACTCACTGAAATGAGCAAAACCAAGTGCCCTCG 7079
Qy 1141 AACACAACCTAGATCCAAACAGAAAATCCACACAGAGACAGACCAACAGCAACACC 1200
Db 7080 ATAAACAACCTCGCTCCAAACGGGAAATCCACACAGCAAAAGAGCACCAGCAGCAAAAAA 7139
Qy 1201 AACATCATCATGACGACATCAGATATAACAAGCAAAACCCCAACAAATTTCTCCGGAT 1260
Db 7140 GGCCCCGCAACACGGCACAAACAGCAAAATGAGCATTTCCACGATCTCTCCCCCAACC 7199
Qy 1261 TCTAGTCGACAAACCGGCCCTCCTATATACCTTTAGAAAAGAACGATCGATCTCTGGAGG 1320
Db 7200 CCCAGCTCGACTGCACAACATCTTGATATTTTCAAGAAAGGGAAGTATCTCTGGAGG 7259
Qy 1321 GAAGCGACATGTTCCCTTTCTCGATGCGTTAAATGATGCTTCAATGATTTTGACCCA 1380
Db 7260 GAAGCGACATGTTCCCTTTCTCGATGCGTTAAATGATGCTTCAATGATTTTGACCCA 7319
Qy 1381 GTTCCAAATACAAAACAACTCTTCATGAATCCTCTAGTCTGCGGCTCGGCTGAGGAA 1440
Db 7320 GTTCCAAATACAAAACAACTCTTCATGAATCCTCTAGTCTGCGGCTCGGCTGAGGAA 7379
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCTTAATATAATAG 1500
Db 7380 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCTTAATATAATAG 7439
Qy 1501 AACACTGCTACTCTGGAGAAAATGAGATGATTTGTGATGCGAGTTAAGAAATTTGGAGC 1560
Db 7440 AACACTGCTACTCTGGAGAAAATGAGATGATTTGTGATGCGAGTTAAGAAATTTGGAGC 7499
Qy 1561 GTTCAGGAGGATGACTGGCCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
Db 7500 GTTCAGGAGGATGACTGGCCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 7559
Qy 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAATTTGCTGCGAGTTGAGG 1680
Db 7560 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAATTTGCTGCGAGTTGAGG 7619
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Db 7620 CGTCTAGCAATCAAACTGCCAAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 7679
Qy 1741 AGAACATCTCCTTAATCAATAGACATGCTATTGATCTTCTACTCACAAGATGGGGAGA 1800
Db 7680 AGAACATCTCCTTAATCAATAGACATGCTATTGATCTTCTACTCACAAGATGGGGAGA 7739
Qy 1801 ACATGCAAAAGTCTTTGGACCTGATTTGTCATCGGATAGAGACTTTGTCAAAAATATT 1860
Db 7740 ACATGCAAAAGTCTTTGGACCTGATTTGTCATCGGATAGAGACTTTGTCAAAAATATT 7799
Qy 1861 TCAGAGCAAAATGACCAATTAATAAGGACGCAAAAAAGAGGAGCTGTTGGGGCTG 1920
|||||

Db 7800 TCAGAGCAAAATTGACCCAAAATTAAAAAGACGACAAAAGAGGAGCTGTTGGGCTG 7859
Qy 1921 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTGCTACTA 1980
Db 7860 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTGCTACTA 7919
Qy 1981 TTATCCATAGTGTCTTGAATGCTCTATCTCTGATATTTGTCGTATCTTTACTAAATATATC 2040
Db 7920 TTATCCATAGTGTCTTGAATGCTCTATCTCTGATATTTGTCGTATCTTTACTAAATATATC 7979
Qy 2041 GGATAA 2046
Db 7980 GGATAA 7985
MREPCYC 19104 bp DNA linear VRL 19-SEP-1996
LOCUS Marburg virus genes for vp35, vp40, vp30, vp24, glycoprotein,
DEFINITION nucleoprotein, polymerase.
ACCESSION Z12132 S55429
VERSION Z12132.1 GI:541780
KEYWORDS glycoprotein; nucleoprotein; polymerase; vp24 protein; vp30
protein; vp35 protein; vp40 protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
REFERENCE 1 Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
2 Sanchez, A., Kiley, M.P., Klenk, H.D. and Feldmann, H.
Sequence analysis of the Marburg virus nucleoprotein gene:
comparison to Ebola virus and other non-segmented negative-strand
RNA viruses
J. Gen. Virol. 73 (Pt 2), 347-357 (1992)
JOURNAL 92166742
MEDLINE 1536192
PUBMED
AUTHORS Muhlberger, E., Sanchez, A., Randolph, A., Will, C., Kiley, M.P.,
Klenk, H.D. and Feldmann, H.
TITLE The nucleotide sequence of the L gene of Marburg virus, a
filovirus: homologies with paramyxoviruses and rhabdoviruses
Virology 187 (2), 534-547 (1992)
JOURNAL 92188528
MEDLINE 1546452
PUBMED
AUTHORS Feldmann, H., Muhlberger, E., Randolph, A., Will, C., Kiley, M.P.,
Sanchez, A. and Klenk, H.D.
TITLE Marburg virus, a filovirus: messenger RNAs, gene order, and
regulatory elements of the replication cycle
Virus Res. 24 (1), 1-19 (1992)
JOURNAL 92327834
MEDLINE 1626422
PUBMED
REMARK 4 (bases 1 to 19104)
AUTHORS Will, C., Muhlberger, E., Linder, D., Slenczka, W., Klenk, H.D. and
Feldmann, H.
TITLE Marburg virus gene 4 encodes the virion membrane protein, a type I
transmembrane glycoprotein
J. Virol. 67 (3), 1203-1210 (1993)
JOURNAL 93172334
MEDLINE 8437211
PUBMED
AUTHORS Feldmann, H.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1992) Feldmann H., Zentrum fuer Hygiene der
Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str.
17, Marburg, Germany, 3550
REMARK 6 (bases 1 to 19104)
AUTHORS Feldmann, H.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-1994) Feldmann H., Zentrum fuer Hygiene der
Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str.

[illegible]

Query Match

74.0%; Score 1514.8; DB 14; Length 19104;

Best Local Similarity 83.8%; Pred. No. 0; Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;			
QY	1	ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT	60
Db	5940	ATGAAGACCATATGTTTCTGATTAGTCTTATCTTATCTTAAATCAAGGACAAAATCTCC	5999
QY	61	GTTTTAAAGATTTGCTAGTAAGACGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAAC	120
Db	6000	ATTTTAAAGATAGCTAGTAATAATCAACCCAAATGTGGATTGGTATGCTCCGGAAT	6059
QY	121	CTCCAAAGACAGAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAGTGTCT	180
Db	6060	CTCCAAAGACAGAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAGTGTCT	6119
QY	181	GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	240
Db	6120	GATTCCTCTTTGGAAGCATCTAAAGCATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	6179
QY	241	GTGAGTATACGGAAGGAGAGAACCCAAACATGTTACAATAATAGTGTAAACAGACCT	300
Db	6180	GTGAGTATACGGAAGGAGAGAACCCAAACATGTTACAATAATAGTGTAAACAGACCT	6239
QY	301	TCTGGAATCTCTGCTGTGATCTCTCCAGTAAATATCCGGATTAACCTTAATGTAAA	360
Db	6240	TCTGGAATCTCTGCTGTGATCTCTCCAGTAAATATCCGGATTAACCTTAATGTAAA	6299
QY	361	ACTGTTTCATATATCAAGGTCAAAACCTCATGACAGGGGATGCTCCCTCATTTTGG	420
Db	6300	ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGGATGCTCCCTCATTTTGG	6359
QY	421	GGGGCATTTTCTGTATGATCGCTGCTCTACAACAATGTACCGAGGCAAGTCTTC	480
Db	6360	GGAGCATTTTCTGTATGATCGCATTGCTTCCACAACAATGTACCGAGGCAAGTCTTC	6419
QY	481	ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAAGAATGATTTTCT	540
Db	6420	ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAAGAATGATTTTCT	6479
QY	541	AGGCAAGGACAGGTTATCGTCACATGAATGACCTCCACCAATAAATATTTGGACAGC	600
Db	6480	CGGCAAGGACAGGTTATCGTCATATGAATCTGACTTCTAATAAGACAGTGCACAAAATGATTTCTCG	6539
QY	601	AGCAATGAAACCGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC	660
Db	6540	AGTAACGGNACCGAACGNAATGACATGATGTTTCGGGCTCTTCAAGAATACAACTCT	6599
QY	661	ACAAACAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG	720
Db	6600	ACAAAGAACCAAAATGCTGCTCCGTCCAAATAATACCTCCACCACTGCCACAGCCGCTCG	6659
QY	721	AGCATTCACTCTCAAAATACCAATTAATTAATCTGCTGAATCTGGAACCTATGACCCAACT	780
Db	6660	GAGATCAAACTCAAGACACCCCAATGATGCCAACAACTCAATACCAAGGACCCAACT	6719
QY	781	AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Db	6720	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGCTCGGAGAACAGAACCCACACA	6779
QY	841	ACTCTTAATGATGATCACTGAACAGAAACAAATGCTCAACAATATTTGTCACCTCTTCACTA	900
Db	6780	ACTTCTGATGCGGTCAACAAAGCAAGGGCTTTTCATCAACAATGTCACCACTCCCTCACTA	6839
QY	901	CATCCAGCACCTCAACATGAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAACT	960
Db	6840	CAACCAAGCACCCCAAGAGGAAACAAACAACCACTATCCCAAGATGCTGTGACT	6899
QY	961	GAGCAATGGAACCGACCCCAACAAACAGCAACAGCAAGCTCTCTCAACAATACTAATACA	1020
Db	6900	GAACTAGACAAAATAACACAACTGCACACCGTCCATGCGCCCTCATAACTACCACTA	6959
QY	1021	ACTCCACCTTAAACATCTCTAAGTACAACTCTGATCTCTCTTCCCTCCACCCGGAAC	1080

RESULT 6
AY430366


```
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGTGATGCAGAGTTAAGAATTTGGAGC 1560
Db 7441 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGTGATGCAGAGTTAAGAATTTGGAGC 7500
QY 1561 GTTCAGGAGGATGACTCTGGCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGAAAT 1620
Db 7501 GTTCAGGAGGATGACTCTGGCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGAAAT 7560
QY 1621 GAAGGACCTTTACACTGCTGTTTTTAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 1680
Db 7561 GAAGGACCTTTACACTGCTGTTTTTAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 7620
QY 1681 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 7621 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 7680
QY 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTACAAAGATGGGAGGA 1800
Db 7681 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTACAAAGATGGGAGGA 7740
QY 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTGTCCAAAAATATT 1860
Db 7741 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTGTCCAAAAATATT 7800
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACCAACAAAAGAGGGAGCTGTTGGGCTCTG 1920
Db 7801 TCAGAGCAAAATGACCAAAATTAATAAGGACCAACAAAAGAGGGAGCTGTTGGGCTCTG 7860
QY 1921 GGTGTAATGCTGGACATCGGACTGGGCTGTTCTTAACTTGGGCAATTTGCTACTA 1980
Db 7861 GGTGTAATGCTGGACATCGGACTGGGCTGTTCTTAACTTGGGCAATTTGCTACTA 7920
QY 1981 TTATCCATAGTCTTGAATGCTCTATCTGATTTGCTGATTTGCTATCTTAAATATATC 2040
Db 7921 TTATCCATAGTCTTGAATGCTCTATCTGATTTGCTGATTTGCTATCTTAAATATATC 7980
QY 2041 GGATAA 2046
Db 7981 GGATAA 7986

RESULT 7
AY430365 19113 bp RNA linear VRL 06-FEB-2004
LOCUS Lake Victoria marburgvirus strain pp3 guinea pig lethal variant,
DEFINITION complete genome.
ACCESSION AY430365
VERSION AY430365.1 GI:40388387
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filooviridae; Marburg-like viruses.
1 (bases 1 to 19113)
Hevey,M., Negley,D., Geisbert,J., Jahrling,P. and Schmaljohn,A.
Antigenicity and vaccine potential of Marburg virus glycoprotein
expressed by baculovirus recombinants
Virology 239 (1), 206-216 (1997)
98087840
MEDLINE
PUBMED 9426460
REFERENCE 2 (bases 1 to 19113)
Chain,P.S.G., Malfatti,S.A., Hajjaj,A., Vergez,L.M., Do,L.H.,
Smith,K.L. and McCready,P.M.
Direct Submission
Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore
National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
3 (bases 1 to 19113)
Ichou,M.A., Paragas,J., Jahrling,P.B., Ibrahim,M.S., Lofts,L.,
Hevey,M. and Schmaljohn,A.
Direct Submission
Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort
Detrick, MD 21702, USA
Location/Qualifiers
```

source

1. 19113

```
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="pp3 guinea pig lethal variant"
/isolate="Museke Kenya, 1980"
/db_xref="taxon:11269"
```

gene

104..2191

```
/gene="NP"
/notes="NP"
/codon_start=1
/product="NP protein"
/protein_id="AAR85460.1"
/db_xref="GI:40388388"
```

CDS

104..2191

```
/gene="NP"
/notes="NP"
/codon_start=1
/product="NP protein"
/protein_id="AAR85460.1"
/db_xref="GI:40388388"
/translation="MOLLSLELGTPTAPHVRNKKVILPDTNHQVSIQNIIDAINS
GIDGLILEGLLTLCVEHYNSDKOKFNTSPIAKYLRDAGYEDVTKNADATFLDV
IPNPHYSPLILALKLTLESTESQGRIGLFLSCSLFLPKLVGDGRASIEKALRQVTV
HORGIVTYPNHMLTTGHMKVIIGILRSSFLKFLVLIHQGNLVLTGHDAYSITNSV
GQTRFSGLLIVKTVLEFILKQTSQGVTLHLFVTSKVNVEASFQKALSILARGEYA
PFARVNLGINNLEHGLYPLSAIALGVATAHGSTLAGVNVGQYQOLREAAHDAEV
KLORHEHQIQAIEDDERKILEQPHLQTEITHSOTLAVLSQREKILARLAABIE
NNTVEDGFKQSNRVSQSFINDPTVEVTQARPMNRPTALPPVDDKIEHSTEDS
SSSSFDVLDNDFALLNDEDETLDDSVMI PGTTSRBFQGIPEPRQSDLNNSQKOB
DESTNPLKKQFLYQELPYOEDDESVTDDSOESIDQSGDNEQGVLDLPPPLXAOE
KRQDTPQHPANFQDFPGSIGDVNGDILEPIRSPSSSAPOEDTRMREAVELSPDFTN
DEBNQQNPQORVVTTKGRTFLYPNDLLQTNPPESLIITALVEEYQNPVSAKELQADWPD
MSFDERRHVAMNL"
```

gene

2945..3934

```
/gene="VP35"
/notes="VP35"
/notes="minor component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="VP35 protein"
/protein_id="AAR85461.1"
/db_xref="GI:40388389"
```

CDS

2945..3934

```
/gene="VP35"
/notes="VP35"
/codon_start=1
/product="VP35 protein"
/protein_id="AAR85461.1"
/db_xref="GI:40388389"
/translation="MMDSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRKPKGTGVS
LQCSPLMSKATSDDDIWDQLVKTLDLIPINQIISDIQSTLSEVTRVHEIER
QJHEITFVLKMGRTLEAISKGSEMLAKYDHLVISTGRTTAPAAFDAYLNEHGVPPP
QPAIFKDLGVAQQAQCSKGTVMKNATDAADKMSKVLSEBETFSKPNLSAKDLALLF
THLPGNNTPFHILAQVLSKIAVKGSGGAFDAPHOILSEGENAQAALTRLSTRFDAP
LGVPVPIRVKNFQTPRCPQKSLRAVPPNPTIDKGWVCYSSSQEGETRALKI"
```

gene

4568..5479

```
/gene="VP40"
/notes="VP40"
/codon_start=1
/product="VP40 protein"
/protein_id="AAR85462.1"
/db_xref="GI:40388390"
```

CDS

4568..5479

```
/gene="VP40"
/notes="VP40"
/codon_start=1
/product="VP40 protein"
/protein_id="AAR85462.1"
/db_xref="GI:40388390"
/translation="MASSSNVNTYMOYLNPPYADHGANOLI PADOLSNQOQITPNYV
GDNLDDQFKNGVCHAFTLEALIDISAYNERTVKGVPALPLGIMSNFSEYELAHTVAA
LLTGSYITQFTGTHQKXFPVNRNLCTGIPAPHLBMLREGNQAFQINWVIRKFNSTNQF
TYNLNLVLSVQKLPDADWPSKNGKLGNTMHPAVSIHPNLPIVLPTVRKQARQKH
NPNNGPLAISGILHLQLRVEKVPKTSLSRISLPADMFVKEGMKKRGNSPVYFQ
APENFPLNGFNRRQVLAAYANPTLSAV"
```

gene

5941..7986

```
/gene="Gp"
/notes="Gp"
/codon_start=1
/product="Gp protein"
/protein_id="AAR85463.1"
/db_xref="GI:40388391"
```

CDS

5941..7986

```
/gene="Gp"
/notes="envelope transmembrane glycoprotein"
/codon_start=1
/product="Gp protein"
/protein_id="AAR85463.1"
/db_xref="GI:40388391"
/translation="MKTCTFLISILIQGTNKLPILEIAGNNQPNQVDSVCSGTLQKLT
EDVHLMGFTLSGQKVAQSPLEAKRWAPRTGVPKPNVEYTEGEBAKCTYNIISVDPDG
KSLILDPTNIRDYPKCKTIHHIQONPHQAQIALHLWGAFFLYDRLASTTMYRGKQVE
TEGNIAMIYKTVHKMIFSRQOGYRHMNLTSTNKYMTSSNGTQNTDNCFCGALOEY
NSTKNQTCAPSKIPPLPLTARPEIKLTSTPTDATKLTATDPSSDELDLAYSGSGSER
EPHTTSDAVTKQGLUSTMPPTFSPQSTPOGGNNTHSQDAVTELDKNNTAQPSMP
```

```
PHNTTITSTNNTSKHNFSTLSAPLQNTTNDNTQSTITENBQTSAPSITTLTPTGNPTT
AKSTSSKGPATTPANTNFTSPPTPTTAQHLVYFRKRSLILRMGDMFPFLDG
LINAPIDPDPVNTKTIIDESSSSGASAEEDQHASPNLSLSTYFPNINEMTAISGEN
ENDCAELURISWVDEEDTLAAGLSMIPPPGCGIEGLYTAGLTKNONNIVCLRRIANOT
AKSLLELLVETTESLALHAIIDFLITWGTGCKVLGPDCCIGIDLSICRITKIYG"
DQIKDEQEGTGWGLGKNWISDMGVLTNLGILLLSIAVLIALLSICRITKIYG"
8869..9714
/genes="VP30"
CDS
8869..9714
/genes="VP30"
/codon_start=1
/product="VP30 protein"
/protein_id="AAR85464.1"
/db_xref="GI:40388392"
/translation="MQQPRGRSRTNRHOVPTIYHETOLPSKPHYTNHPRAGMSST
RSSAESPTNHPRARPPSTFNLSKPPPPKDMCRNMKIGLPCADPTCNRDHDIDNLT
NRELLILMARKNLNTDKTFRSPDCGSPSLSKGLSKDKQKQOTKDVLTLENLHLSY
LHRSIGLKDTELRAALUSLTACAGIRKTRSLNTLHNMHNHNLQDQNGVJKQTY
TGIHLDDGGQFEAALWQGDKRISLSLFOAALYVMNNIPCESSISVQASDYHFLPOS
QGGQ"
10207..10968
/genes="VP24"
CDS
10207..10968
/genes="VP24"
/codon_start=1
/product="VP24 protein"
/protein_id="AAR85465.1"
/db_xref="GI:40388393"
/translation="MAELSTRYNLPANTYENSINLDLNSTARWIKERSVSGWTVKGN
FVFHPNTGMILLHLKSNFVVPQQTRNLFSLFKNPKSTIIEPFLALRILGLVAL
KDQELQOQLIFGFSIYHMLSEWLLLEVTSAIHSNLLGIYLSIDMFKILMAGVKNP
FNKFTLHVNDHGKSPSIEIKLQGOQIIITRVNMGFLVEVRRIDIEPCCGTVLSES
VVFGLVABAVLRHSQMEKGQPLNLTYMNSKIAI"
11481..18476
/genes="L"
CDS
11481..18476
/genes="L"
/notes="major component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="L protein"
/protein_id="AAR85466.1"
/db_xref="GI:40388394"
/translation="MQHPTQYDPARLSPIILDQDLRLSLGLSYSHVSHNPKLRNCR
IPHILYRLNRSTALKTFLQNGSILTPFHSIWDHILTSIQYDAINHVDFFKYLPSL
VKYANWDEFKAYLANKLGLDHPVASARQCEDFSPKNPYWGLMLLVHLSQLAR
RIKQSGSLRNKWFKIGTDLRGLADFIKVPVKTIIRNAVSLQASKPGLRIYWRD
QNLTFLCDDDEFIVSAVECFIMKDVIERYNWTEICARAWLEDSGADYPLDVL
GELYNQGOQITAMYLEDDGFKLIKHLEPLVCVCIQTHGIFTPRKYWFQSQMKSYYDEL
HDNLKQISDNKAECAQNFIKTIQAKLTPOQCELSLQKHGHPVLYNDVALDKV
KHAQSTKILKPKVMFVFCVKFIKAKNHYSQGSWYKTHDLHLTPYLRQHIVSNS
PFSQAEIYQHLWEWYFVBEHPLFTKIIISLSIPIKORATAVNOECWDSVFDRLVGY
NPPVRFQSKRPYQFLGQADFSLNQILFEAEKLEVLAPSFRNFSLSKEKELNIGRTP
GKLPYRVNQTLEALLADGLAKAPPSNMVMVTEREQEALLHQASHHNSHISGEN
AIVRGASFVTDLEKTLNLAIFYETKHFIDYCNRCYGVKNLFDMMHFLIPLCYMVSDF
YSPPHCVTDRNRNPPDCANAYHYHLGGIEGLQKQKLTCSICAQITVELTKKLKLS
CVMGDNQCTITLSLPPIDAPNDYQNEAELNAARVAVELAITTGYSGIFLKPEFTVH
SGFIYFGKQYVNLQPSLKTWARCPLSDSIDDLQGLSAGTISFPERGTSTHGH
IFPSWIASFMSLAINLQNLHGLFPFGENDIDSCFKPLTFSEKILALITPOVLGH
LSFLNPEKLFYRNISDPLTSLGFLQKNALFELEKEELFYIILSKPGLADASDFWNP
LGLNVPGSKELITFURQVRENITISQNRILINSFLHIGSDLEQDORVCEWLLSNPVN
SRFAADITSRTSPGKQVLTGLETRTLASRTISLTETGLMKRLMRLTRNWKWS
FSYIDALDDLSLESLEKFTCTVDVANFLRAYSWSDVLKGRLLIGATPLCLIEQEVKW
INLSDELQOENLSDSKSTINLLPYDCKELRLEGSNDTELNYVSCALDRKVVQKHS
VNRLAWTIGNRAPYIGSRTEDEKIGYPPLRVNCPSAALKEATEMWSRLLLWVTQTADRE
KLLIFLLSNVNDLYQSTENKELPHYSGNIVHRYNDQYQHSFMANRNSVSTRALIS
TNLKGYAGGQAAIDSNILIFQNTINLGVAVLDIALSLKSSASNSVTFRMLNKCCT
RHVPSEYIYFPKPLDVLKYNMDELVDNDPLCSGKIGRIGRVSRSTLSLNSVSDI
GSYDPTTAAMTGLSTIVGSIIFSDESSQSTDPISGGCTKTFVTHFVPSVSIIFYAG
ANLIYVESLSLRKSIKNSLDTPLISSTIRNLHRSRLILQSTFRHELPLTALHAI
PLISLMGSGAGEKSSDAPRLFLTASYNFINNPFCLMKKQSLSPLWLYPSEGOQ
LKPILKILQRLSDLSLPKIKQRKILADTCCPIGFSFWYPSKSTRNHYASLNYWRD
KANKVNTFFSHLINCSPFESSSTSVSSNQVTSNKIIVIPENITBAINARTLLINY
GSTALQGMDDTKMPLNSEQNLENCRPSESGIRFDKNQKIKTHDQRCEREBESSPQQMFED
```

```

NMQTPAHIHSSSPFOILLKSLDAHEDFDASKIILNSEINNINLTYLTNTKLLTTPTR
TEILDTPSASRYSTSRSSRLSREQASYLYVDCSNIPISIDPFRSPSSDQNOVQ
MLINTYKRDHLACFDNQFRFTQSSSMHYKLYDLLPPGKLKAIICLAIEGSGGARL
LLKWKEDYLFENFTLATDSQEBAILSGRIVPRMLYNIIDRLSALLSRRLLLNLTIQ
ITDITNPLMDSVIQVLPESDILTMDAETTKDREQLVKTIVNIWRTSPNIPKIS
IIKVLDDYEGTILFMKNAIQYGVQVQKPKPYSSNNAKSEWYLCGRRRIORLQIDFS
DQVGFILICAMSRQRAIPIWHLKHIEKNYPASLHEFFLTIGFSPLESSEFCHRYTIPF
SEKALFHKVQSYVRQKQHLHSLMDYENNSPLLDLRNHFICSLRGKITKYNDILK
LNLVIAKEGKNWSQVLEILPNMHSVCIVHVDHECSGCEKRLLLKLDIFIRNTKIAEQ
KLLNRVIGYILFPFPGLPKSGSLRA"
ORIGIN
Query Match 74.0%; Score 1513.2; DB 14; Length 19113;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Dd 5941 ATGAAGACCATATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC 6000
Qy 61 GTTTTATAGAAATGCTAGTAACAGCAACCTCAAGATGTAGATTAGTGTGCTCCGGAAACC 120
Dd 6001 ATTTTATAGATAGTAGTATATAATCAACCCAAAATGTGGATTCGTATGCTCCGGAAT 6060
Qy 121 CTCCAAAGACAGAAGATGTTTCATCTGATGGGATTTACACGTAGTGGGCAAAAAGTTGCT 180
Dd 6061 CTCAGAAGACAGAAGACGTCATCTGATGGGATTTACACGTAGTGGGCAAAAAGTTGCT 6120
Qy 181 GATTCCTCTTTGGAGCATCTTAACGATGGGCTTTTCAGGACAGGTGTTCCCTCCCAAGAAC 240
Dd 6121 GATTCCTCTTTGGAGGATCAACGATGGGCTTTTCAGGACAGGTGTTCCCTCCCAAGAAC 6180
Qy 241 GTTCAGTATACGGAAGAGAGAAGCAAAACATGTTACAATATAAGTGTAAACAGACCT 300
Dd 6181 GTTCAGTACACAGAGGGGGAGGAAGCCAAAACATGCTACATATAAGTGTAAACAGATCC 6240
Qy 301 TCTCGAAAAATCTTTGCTGTGGATCTCTCCAGTAAATATCCCGATATACCTTAATGTAAA 360
Dd 6241 TCTCGAAAAATCCTTGCTGTAGATCTCTTACCAACATCCGTGACTATCTTAATGCAAA 6300
Qy 361 ACTGTTTCATCATATTTCAAGTCAAAACCTCATCGACAGGGGATTCCTCCATTTGTTGG 420
Dd 6301 ACTATCCATCATATTTCAAGGTCAAAAACCTCATCGACAGGGGATTCCTCCATTTATGG 6360
Qy 421 GGGGCATTTTCTTGATGATCGCGTGTGCTCTACAACAATGTACCGAGCAAGTCTTC 480
Dd 6361 GGAGCATTTTCTTGATGATCGCATTTGCTCTCACACAATGTACCGAGCAAGTCTTC 6420
Qy 481 ACTGAAGAAATATAGCAGCTATGATTGTTAATAGACAGTTCACAGATGATTTTTTCT 540
Dd 6421 ACTGAAGGAACATAGCAGCTATGATTGTTCAATAAGACAGTGCACAAAATGATTTCTCG 6480
Qy 541 AGGCAAGGACAGGTTATCGTCACATGAACTTGACCTCCACCAATATAATTTGGACAAGC 600
Dd 6481 CGGCAAGGACAGGAGTACCGGTACATGAACTCTGACTTCTACTAATAAATATTTGCAAGT 6540
Qy 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGAAATACAATCC 660
Dd 6541 AGTAACCGNAGCCAAAACGAATGACATGGAATGTTTGGCGCTCTTCAAGNATACAATTTCT 6600
Qy 661 ACAAAATCAAAACATGCCCTCCATCTCTTTAAACCTCCATCCCTGCGCCACAGTAAGTCCG 720
Dd 6601 ACAAGAAACCAACATGTGCTCCGTCCAAAATAGCTCTCCACCACTGCCCCACAGCCGTCG 6660
Qy 721 AGCATTCACCTACAAATACTCAAAATTAATACTGTAATCTGGAACCTATGAACCAAGT 780
Dd 6661 GAGATCAAACTCAACAGCACCCCAACTGATGCCCAAACTCAATACCCAGGACCAAGC 6720
Qy 781 AGCGACGATAGGACCTTATGATTTCCGGCTCAGGATCTGAGAAACAGGGGCCCCACACACA 840
Dd 6721 AGTGATGATAGGACCTCGCAACATCCGGCTCAGGGTCCGAGGAACGAGAACCCACACACA 6780
Qy 841 ACTCTTAATGTAGTCACTGAAAGAAACAAATCGTCAACAATATTGTCTCACTCTTCACTA 900
```

Db 6781 ACTTCGTGATGGGTGACCAAGAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACCA 6840
Qy CATCAAGCACCTCACAACATGAGCAAAAGTAGTACGAATCTCTCCGACATGCTGTAACT 960
Db 6841 CAACAAGCAGCCACAGCAGAGGAGGAACAACAACCAATTTCCCAAGATGCTGTGACT 6900
Qy 961 GAGCAAAATGGAACGAGCCCAACAACAACAACAGCAAGCTCTCTCAACAATACTAATAACA 1020
Db 6901 GAACTTGACAAAAATAACAACAATGCAACAACGCTCCATGCGCCCTCATAAACATACCACA 6960
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 6961 ATCTTCTAACAACACTCTCAACAACACTTTCAGCACTCTCTCTGCAACCAATTAACAAC 7020
Qy 1081 ATCAACCAATATGATACACAACAGTGAATAGCAGAAAGCGAACAACCAATGCTCAGTTG 1140
Db 7021 ACCACCAATGACAACAACAGCAGCAATCACTGAATATGAGCAACCCAGTGCCCTCG 7080
Qy 1141 AACCAACTCTAGATCCAAAGAAAATCCCAACAGCAGCAAGACCAACAAGCAGCAAC 1200
Db 7081 ATAACAACCTGCTCCCAACGGAATCCCAACAGCAAGACCAACAGCAGCAAAAA 7140
Qy 1201 AACATCATGACGACATCAGATATACAGCAACCAACCAACCAATTTCTTCCGGAT 1260
Db 7141 GGCCCGCCCAACGCGCAACCAACAGCAATGAGCATTTACAGTCTCCCTCCCAAC 7200
Qy 1261 TCTAGTCGCAACCGCGCTCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 7201 CCCAGCTGACGACACACACTCTGTATATTTAGAGAAAGCGAGTATCTCTGGAGG 7260
Qy 1321 GAAGCGCATGTTCCCTTTCTGATGGGTAAATAAATGCTCCAAATGATTTGACCCA 1380
Db 7261 GAAGCGCATGTTCCCTTTCTGATGGGTAAATAAATGCTCCAAATGATTTGACCCA 7320
Qy 1381 GTTCCAAATACAAAACAACTCTTTGATGAATCTCTAGTTCTGGTCTCGCTGAGGAA 1440
Db 7321 GTTCCAAATACAAAACAACTCTTTGATGAATCTCTAGTTCTGGTCTCGCTGAGGAA 7380
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAAATAAATAAGAG 1500
Db 7381 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAAATAAATAAGAG 7440
Qy 1501 AACCTGCTACTCTGGAGAAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTGGAGC 1560
Db 7441 AACCTGCTACTCTGGAGAAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTGGAGC 7500
Qy 1561 GTTCAGGAGATGACTCGCGCAGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAAT 1620
Db 7501 GTTCAGGAGATGACTCGCGCAGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAAT 7560
Qy 1621 GAAGGACTTTACACTGCTGTTTAAATTAATAATCAAAAATAATTTGGTCTGCAGGTTGAGG 1680
Db 7561 GAAGGACTTTACACTGCTGTTTAAATTAATAATCAAAAATAATTTGGTCTGCAGGTTGAGG 7620
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCTCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 7621 CGTCTAGCAATCAAACTGCCAAATCTCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 7680
Qy 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGATGCTTTCTACTCAAGAATGGGAGGA 1800
Db 7681 AGAACATTTCTCTTAATCAATAGACATGCTATTGATGCTTTCTACTCAAGAATGGGAGGA 7740
Qy 1801 ACATCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTCTGTCCAAAATAAT 1860
Db 7741 ACATCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTCTGTCCAAAATAAT 7800
Qy 1861 TCAGAGCAAAATGACCAAAATTAAGGACGCAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 7801 TCAGAGCAAAATGACCAAAATTAAGGACGCAACAAAAGAGGGGACTGGTTGGGGCTG 7860
Qy 1921 GGTGTAAATGTTGGACATCCGACTGGGGTGTCTTAATACTTGGGCAATTTTGTCTACTA 1980

Db 7861 GGTGGTAATGGTGGACATCCGACTGGGTGTTCTTACTAACTTGGGCAATTTTGTCTACTA 7920
Qy 1981 TTATCCATAGCTGTCTTGATGCTCTATCTGATTTGTCGTATCTTTTACTAATAATATC 2040
Db 7921 TTATCCATAGCTGTCTTGATGCTCTATCTGATTTGTCGTATCTTTTACTAATAATATC 7980
Qy 2041 GGATAA 2046
Db 7981 GGATAA 7986
RESULT 8
AX717721
LOCUS AX717721 2051 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 9 from Patent WO02079239.
ACCESSION AX717721
VERSION AX717721.1 GI:29890729
KEYWORDS
SOURCE Marburg virus (strain Musoke)
ORGANISM Marburg virus (strain Musoke)
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1
AUTHORS Grogan, C. C., Hevey, M. C. and Schmaljohn, A. L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 9 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
LOCATION/Qualifiers
FEATURES
1..2051
source
/organism="Marburg virus (strain Musoke)"
/mol_type="unassigned DNA"
/db_xref="taxon:33727"
/note="chimeric molecule between Marburg virus
Glycoprotein 1 and Marburg virus Glycoprotein 2"
ORIGIN
Query Match 72.6%; Score 1485.6; DB 6; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAATCCAAAGTATAAAAACCTCTCCCT 60
Db 10 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAATCTCCCC 69
Qy 61 GTTTTAGAAATGTCTAGTAACAGCAACCTCAAGATGTAGATTCACTGTGTCTCCGGAAC 120
Db 70 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCTGCTATGCTCCGGAAC 129
Qy 121 CTCGAAAGACAGAAGATGTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 130 CTCGAAAGACAGAAGACGCTCCATCTGATGGGATTCACCTGAGTGGGCAAAAAGTTGCT 189
Qy 181 GATTCCCTTTTGGAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 190 GATTCCCTTTTGGAGGATCCCAACGATGGCTTTTCAGGACAGGTGTTACCTCCCAAGAT 249
Qy 241 GTTCAGTATACGGAAGGAGAGCAACCAATGTTTACAAATAGTGTAAAGTGTAAACGACCT 300
Db 250 GTTCAGTATACGAGGGGGAGAGCAACCAATGTTTACAAATAGTGTAAAGTGTAAACGACCT 309
Qy 301 TCTCGAAATCTCTTGTGCTGATGCTCCCGATTAATATCCCGATTAATATCCCTCAATATGTTG 360
Db 310 TCTCGAAATCTCTTGTGCTGATGCTCCCTACCAACATCCCGTACTATCCGAAATGCAAA 369
Qy 361 ACTCTTCATCATATTAAGGTCAAAACCTCATCATGAGGGGATTCCTCATTTGTTGG 420
Db 370 ACTATCCATCATATTAAGGTCAAAACCTCATCATGAGGGGATTCCTCATTTATGTTG 429
Qy 421 GGGGCAATTTTCTTGTATGATCGGTTGCTCTCAACAATATGACCGAGGCAAGGTCTTC 480
Db 430 GGAGCAATTTTCTTGTATGATCGCAAT-----ATGTACCGGAGGCAAGGTCTTC 477
Qy 481 ACTGAAGGAATATAGCAGCTATGATTGTTTAAATAGACAGATTTCACAGATGATTTTTTCT 540

||||| 478 ACTGAAGGAACATAGCAGCTATGATGTCAATAGACAGTGCACAAATGATTTTCG 537
||||| 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTGGACAGC 600
||||| 538 CGGCAAGGACAGGCTCGTCATATGATCTGACTTCTACTATAAATATTGGNCAAGT 597
||||| 601 AGCAATGAACGCGAGAGAAATGATACGGGATGTTTGGGATCTCTCAAGAATATACACTCC 660
||||| 598 AGTACGGAACGCAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT 657
||||| 661 ACAACCAATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
||||| 658 ACAAGAACCAACATGTGCTCCGTCACAAATACCTCCACACTGCCACAGCCCGCTCG 717
||||| 721 AGCATCTCTACAAATTAATACTCAAAATTAATACTGCTAAATCTGGAACCTATGAACCCAAGT 780
||||| 718 GAGATCAAACTCACAAACCCCACTGATGCCCAAACTCAATACACGACGCCAAGC 777
||||| 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAGAACAGGGGCCCCACACA 840
||||| 778 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGGAGAACGAGAACCCCAACA 837
||||| 841 ACTCTTAATGTAGTCACTGACAGAAACAACTCGTCAACAATATGTCCACTCCTTCACTA 900
||||| 838 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 897
||||| 901 CATCCAAGCAGCTCACAAATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
||||| 898 CAACCAAGCAGCCACAGGAGGAGAAACAACAACCAATCCCAAGATGCTGTGACT 957
||||| 961 GAGCAATATGAAACGACCCCAACAACAACAGCAACGCTCTCTCAACAATACTAATACA 1020
||||| 958 GAACTAGACAAAATAACAACTGCAACCGTCCATGCCCTCTATAAACAATACTACCACA 1017
||||| 1021 ACTCCCACTTAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGGCAAC 1080
||||| 1018 ATCTCTACTAAACACCTCCAAACACAACTTCAGCAGCTCTCTGTGCACCAATTAACAAAC 1077
||||| 1081 ATCACCAATATGATACAAACGTAAGTACGAGAAAGGCAACCAACCAATGCTCAGTTG 1140
||||| 1078 ACCACCAATGACACACAGCAGCAATCACTGAAATGAGCAACCAAGTGCCTCG 1137
||||| 1141 AACACAATCTAGATCCCAAGAAAATCCCAACAGGACAGCAACCAACAGCAACCAAC 1200
||||| 1138 ATAAACACCTGCTCCCAAGGAAAATCCCAACAGCAAGAGCAACCAAGCAAGCAAAA 1197
||||| 1201 AACATCATGTACGACATCAGATATACAGCAACACCCCAACAAATCTTCTCCGAT 1260
||||| 1198 GGGCCCGCCCAACGGCACCAACACGCAAAATGAGCATTTCCACAGTCTCCGCCACC 1257
||||| 1261 TCTAGTCCGACAAACCGCCCTCTATATATCTTGAAGAAACGATCGATCTCTGGAGG 1320
||||| 1258 CCAGCTCGACTGCAACATCTTGATATTTTCAGAGAAAGCGATCGATCTCTGGAGG 1317
||||| 1321 GAAGCGCATGTTCCCTTTCTGGATGGGTTAAATAATGCTCCAAATGATTTGACCCA 1380
||||| 1318 GAAGCGCATGTTCCCTTTCTGGATGGGTTAAATAATGCTCCAAATGATTTGACCCA 1377
||||| 1381 GTTCAAAATACAAAACAACTCTTGAATGATCTCTAGTTCTGTGGTCCCTCGCTGAGGAA 1440
||||| 1378 GTTCAAAATACAAAACAACTCTTGAATGATCTCTAGTTCTGTGGTCCCTCGCTGAGGAA 1437
||||| 1441 GATCAACATGCTCCCAATATAGTTAGTTTAACTTTATCTTATTTTCCCTAAATAAATGAG 1500
||||| 1438 GATCAACATGCTCCCAATATAGTTAGTTTAACTTTATCTTATTTTCCCTAAATAAATGAG 1497
||||| 1501 AACATGCTCTCTGGAGAAAATGAGATGATGTGATGACAGGTTAAGAAATTTGGAGC 1560
||||| 1498 AACATGCTCTCTGGAGAAAATGAGATGATGTGATGACAGGTTAAGAAATTTGGAGC 1557
||||| 1561 GTTCAGGAGGATGACTGGCCGACGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 1620
|||||

Db 1558 GTTCAGGAGGATGACCTGGCCGACAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 1617
Qy 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1680
Db 1618 GAAGGACTTTACACTGCTGTTTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1677
Qy 1681 CGTCTAGCCCAATCAAACTGCCCAATCTTGGAACTCTTATTTGAGAGTCAACAACCTGAGGAA 1740
Db 1678 CGTCTAGCCCAATCAAACTGCCCAATCTTGGAACTCTTATTTGAGAGTCAACAACCTGAGGAA 1737
Qy 1741 AGAACAATTCCTTAAATCAATAGACATGCTATTGACTTTCTACTCAACAAGATGGGGAGGA 1800
Db 1738 AGAACAATTCCTTAAATCAATAGACATGCTATTGACTTTCTACTCAACAAGATGGGGAGGA 1797
Qy 1801 ACATGCAAAAGTGTGTGACCTGATTGTGTGATCGGGATAGAGACTTGTGCCAAAATATT 1860
Db 1798 ACATGCAAAAGTGTGTGACCTGATTGTGTGATCGGGATAGAGACTTGTGCCAAAATATT 1857
Qy 1861 TCAGAGCAAAATTTGACCAATTTAAAAGGACGAAACAAAAGGAGGACTGTGTGGGTCTG 1920
Db 1858 TCAGAGCAAAATTTGACCAATTTAAAAGGACGAAACAAAAGGAGGACTGTGTGGGTCTG 1917
Qy 1921 GGTGTTAAATGTTGGACATCCGACTGGGGTGTCTTACTTAATCTTTGGGCATTTTGTCTACTA 1980
Db 1918 GGTGTTAAATGTTGGACATCCGACTGGGGTGTCTTACTTAATCTTTGGGCATTTTGTCTACTA 1977
Qy 1981 TTATCCATAGCTGCTTTGATTGCTCTATCTCTGATATTTGTCTATCTTTACTAAATATATC 2040
Db 1978 TTATCCATAGCTGCTTTGATTGCTCTATCTCTGATATTTGTCTATCTTTACTAAATATATC 2037
Qy 2041 GGATAA 2046
Db 2038 GGATAA 2043

RESULT 9
AF005735
LOCUS
DEFINITION Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein precursor (GP) gene, complete cds.
ACCESSION AF005735
VERSION AF005735.1 GI:2459879
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Virus; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PubMed 9448638
REFERENCE 2 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA

FEATURES
source
1. .2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Germany/Marburg/1967/Ratayczak"
/db_xref="taxon:11269"
1. .2046
/gene="Gp"
1. .2046
/gene="Gp"
/note="viral surface glycoprotein"
/codon_start=1

```
/product="glycoprotein precursor"  
/protein_id="AAC40460.1"  
/db_xref="GI:2459880"  
/translation="MKTFCLFSLILIQIKTLPILEIASNNQPNVDSVCSGLQKT  
EDVHMGPTLSQKADSPLEBSKRWAFRTGVPKPNVEYTEGEBAKTCYNI SVTDPG  
KSLLDPPNIRDYPKCKTHHIOQONPHAQIALHLWGAPFLYDRISTVMYGRVFE  
TEGNIAMVUNKTVHMKFSQOGGYRHMNLTSTNKYTSNNGTQTNDCGFCGALOEY  
NSTKQTCAPSKISPLPTARPEIKPTSTPTDATTILNTDNDDEBLIISGSGSGQ  
EPTYSOAVTKQGLSSTMPPTSPSPQSPQOEGNNTHDSOGTVTEPKNTINTTAQSPM  
PHNTATSTNNSTKNSTSVLSQNTTNYDTQSTATENBOQTSAPSKTTLPTGNLT  
AKSTNNTGPTTTPAINTNHLTSPSPNPTQHLVYFRKRSILWREGDMFPFLDG  
LINAPIDPVPNTKTTIDESSSGSAEEDOHASPNISLTSLVFPNINENTAYSGN  
ENDCAELRIVSWOEDDLAAGLSMIPFPGIGLYTAGLIGNONNVLCRLRLANOT  
AKSLELLRVTEBTEPSLNRHAIPELLFRWGCTCKVLGPDCCIGIEDLSRNISEQI  
DQIKDKQKQETGWLGGKMWTSIDWGLNLTNLLLSIAVLIALSLCIRIFTKIYG"  
1..81  
/gene="Gp"  
82..2043  
/gene="Gp"  
/product="glycoprotein"  
1930..2019  
/gene="Gp"  
/note="encodes transmembrane anchor peptide"
```

ORIGIN

```
Query Match 72.1%; Score 1474.8; DB 14; Length 2046;  
Best Local Similarity 82.6%; Pred. No. 0;  
Matches 1689; Conservative 0; Mismatches 357; Indels 0; Gaps 0;  
  
QY 1 ATGAAGACCATATATTTCTCATGTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60  
DB 1 ATGAAGACCATATGCTCTTTATCATGCTTATCTTTATCCAAAGGATATAAACTCTCCCT 60  
  
QY 61 GTTTTAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTGAGTGTGTCGCGGAACC 120  
DB 61 ATTTTAGAGATAGCTAGTAACATCAACCCCAAAATGTGATTCGGTATGCTCGGAACT 120  
  
QY 121 CTCMAAAGACAGAAGATGTTCACTGTAGTGGATTTTACATGAGTGGGCAAAAAGTTGCT 180  
DB 121 CTCMAAAGACAGAAGATGTTCACTGTAGTGGATTTTACATGAGTGGGCAAAAAGTTGCT 180  
  
QY 181 GATTTCCCTTTGGAGCATCTAAGAGTGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240  
DB 181 GATTTCCCTTTGGAGCATCTAAGAGTGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240  
  
QY 241 GTTGAGTATAGGAAGAGAGAAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 300  
DB 241 GTTGAGTATACAGAAGGGAGAGCCAAACATGCTACATATATAAGTGTAAACAGATCCC 300  
  
QY 301 TCTGAAAAATCTTGCTGCTGATCCTCCAGTAAATATCCGGAATACCTTAATGTAAA 360  
DB 301 TCTGAAAAATCTTGCTGCTGATCCTCCAGTAAATATCCGGAATACCTTAATGTAAA 360  
  
QY 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGCAAGGGGATTCCTCATTTGTTGG 420  
DB 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGGATTCCTCATTTGTTGG 420  
  
QY 421 GGGCAATTTTCTGTATGATCGCTGCTCTCAACAAATGTACCGAGGCAAGTCTTC 480  
DB 421 GGAGCATTTTCTGTATGATCGCAATGCTCTCAACAAATGTACCGAGGCAAGTCTTC 480  
  
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCACAGAAATGATTTTTCT 540  
DB 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCACAGAAATGATTTTTCT 540  
  
QY 541 AGGCAAGGAAAGTATGTCACATGAACTTGACTCCACCAATATAATATGGACAGC 600  
DB 541 AGGCAAGGAAAGTATGTCACATGAACTTGACTCCACCAATATAATATGGACAGC 600  
  
QY 601 AGCAATGAAACGACAGAAATGATACGGATGTTTGGCATCTCCCAAGAAATACAACTCC 660  
DB 601 AGCAATGAAACGACAGAAATGATACGGATGTTTGGCATCTCCCAAGAAATACAACTCC 660
```

```
Db 1741 AGGACATTTCTTAATTAATAGACATGCCATTGACTTTCTACTCACAAGGTGGGGAGGA 1800
Qy 1801 ACATCAAAAGTGTGGACCTGATGTTGGATCGGATAGAAAGACTTCTGCCAAAATATT 1860
Db 1801 ACATCAAAAGTGTGGACCTGATGTTGGATCGGATAGAAAGACTTCTGCCAGGAATATT 1860
Qy 1861 TCAGACAAATTTGACCAAAATTAATAAGGACGAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 1861 TCGGAACAAATTTGACCAAAATCAAAAAGATGAACAAAAGAGGGGACTGGTTGGGGCTA 1920
Qy 1921 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAATCTGGGCAATTTTGGCTACTA 1980
Db 1921 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAATCTGGGCAATTTTGGCTACTA 1980
Qy 1981 TTATCCATAGCTCTTCTGATGCTCTATCTCTGTTATTTCTGCTGATCTTTTACTAAATATATC 2040
Db 1981 TTATCCATAGCTCTTCTGATGCTCTATCTCTGTTATTTCTGCTGATCTTTTACCAGATATATC 2040
Qy 2041 GGATAA 2046
Db 2041 GGGTAA 2046

RESULT 10
MAVSPAA
LOCUS Marburg Virus genomic RNA of GP gene. linear VRL 22-JAN-1996
DEFINITION
ACCESSION X68493
VERSION X68493.1 GI:296960
KEYWORDS gp gene; structural protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses: ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE
1 (bases 1 to 2948)
Bukreyev, A., Volchkov, V.E., Blinov, V.M. and Netesov, S.V.
The GP-protein of Marburg virus contains the region similar to the
'immunosuppressive domain' of oncogenic retrovirus P15E proteins
PEBS Lett. 323 (1-2), 183-187 (1993)
93265932
8495737
PUBMED
2 (bases 1 to 2948)
Bukreyev, A.A., Volchkov, V.E., Blinov, V.M., Dryga, S.A. and
Netesov, S.V.
The complete nucleotide sequence of the Popp (1967) strain of
Marburg virus; a comparison with the Musoke (1980) strain
Arch. Virol. 140 (9), 1589-1600 (1995)
96028047
7487490
PUBMED
3 (bases 1 to 2948)
Bukreyev, A.A.
Direct Submission
JOURNAL
MEDLINE
REFERENCE
1. 2948
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="Popp"
/db_xref="taxon:11269"
6. 2851
/gene="GP"
(6.17) (2841.2851)
/gene="GP"
122. 2167
/gene="GP"
/codon_start=1
/product="structural protein"
/protein_id="CAA48507.1"
/db_xref="GI:296961"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
```

```
/translation="MKTTCLPISLILQGIKTLPILEIASNNQPNQVDSVCSGLTQKT
EDVHLMFTLSGQKVADSPLEASKRWFAPGVPPKNVYTEGEBAKTCYNTSVDPDSG
KSILDDPTNIRDYPKCKTHHIOQNPHQAQIALHLWGAFFLYDRASTMTYRGRVF
TEGNIAAMVNKTVHKMI FSRGOGYRHMALTSNKYWTNNGTQNDTTCGCGALOEY
NSTKNOFCAPSKIPSPLPARPEIKPTSTPDATLANTDPNNDDEDLITSGSGSEGO
EPTYTDAVTKQGLSMTPTSPQSPQBPQEBNNTHDSQGTVPKNTKNTTAQPSMP
PHNTATSTNNTSKNPFSTLSLQNTNYDTQSTATENEQTSAPSILKTLPTGLNLT
AKSTNTPGTPTTAPNNTNGHLTSPPTNPPTQHLVYFRKRSILNREGDMFPFLDG
LINAPIDFPVPTKTI FDESSSGSAEEDQHASPNISLTLSPNNINENTAYSGEN
ENDCDALRTWSQEDDLAAGLSWIPFGGIEGLYTAGLTKNKNILVCRLLRLANOT
AKSLELLRLRTBERTESLINRHAIDFLLRWGSTCKVLGPDGCTGIEDLSRNTSEQI
DQIKORQKSGTGWLGKRWKTSWGVITNLGILLLSIAVLIALSICRIFTKYIG"

ORIGIN
Query Match 72.0%; Score 1473.2; DB 14; Length 2948;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAATCCAAAGTATATAAACTTCCTCT 60
Db 122 ATGAAGACCACATGCTCTTTATCAGTCTTATCTTAATCCAAAGGATATAAACTTCCTCT 181
Qy 61 GTTTTAGAATTTGCTAGTAAAGCAACCTCAAGATGTAGATTTCAGTGTCTCCGGAACC 120
Db 182 ATTTTAGAGATAGTAGTAACAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 241
Qy 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTTACACTGAGTGGGCAAAAGATTGCT 180
Db 242 CTCAGAAGACAGAAAGATGTCATCTGATGGGATTCACACTGAGTGGGCAAAAGATTGCT 301
Qy 181 GATTCCCTTTTGAAGCATCTAAACGATCGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 302 GATTCCCTTTTGAAGCATCTCAAGCATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 361
Qy 241 GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTACATATAAGTGAACAGACCT 300
Db 362 GTTGAGTATACAGAAGGGAGAGAGCCAAACATGCTACAATATAAGTGAACAGGATCCC 421
Qy 301 TCTGAAAATCTCTTGCTGCTGATCTCCCAAGTAAATATCCCGGATTTACCTAAATGTAAA 360
Db 422 TCTGAAAATCTCTTGCTGCTGATCTCTTACCAACATCCGCTGACTATCTTAATGCAAA 481
Qy 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATTTGCCCTCCATTTGCG 420
Db 482 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAGAGGATCGCCCTCCATTTGCG 541
Qy 421 GGGCATTTTCTGTATGATCGCGTTCCTCTACAACATGTAACCGAGGCAAGGTCTTC 480
Db 542 GGAGCATTTTCTGTATGATCGCATTTGCCCTTCCAACAATGTACCGAGGAGAGATCTTC 601
Qy 481 ACTGAAGGAATATAGCAGCTATGATTGTTTAATAAGACAGTTCACAGAATGATTTTTTCT 540
Db 602 ACTGAAGGAACATAGCAGCTATGATTGTTCAATAAGACAGTTCACAGAATGATTTTTCTCG 661
Qy 541 AGGCAAGGACAAGGTATTCGTCATGAACCTTGACCTCCACCAATAATATTTGGACAAGC 600
Db 662 AGGCAAGGACAAGGATACCGGTCAATGAATCTGACTTCTACTAATAATATTTGGACAAGT 721
Qy 601 AGCAATGAAGGACAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGATATCAACTCC 660
Db 722 AACAAATGGACACAAACGAATGACACTGGATGCTCTTCAGAAATATCAACTCC 781
Qy 661 ACAAAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAAGTCCG 720
Db 782 ACGAAGAAATCAAAACATGTGCTCCGTCCAAAATACCTCACCCTGCCACAGAGCCGCTCA 841
Qy 721 AGCAATTCATCTACAAATACCTCAAAATTAATAACTGCTAAATCTGGAACACTATGAACCAAGT 780
Db 842 GAGATCAAAACCAAGCACCCCAACTGATGCCCAACACACTCAACACCCACAGACCCCAAC 901
Qy 781 AGCCAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGACAGAGGGGCCCCACACA 840
Db 902 AATGATGATGAGGACCTCATAACTCCGGTTTCAGGGTCCGGAGAACAGGAACCCCTTATACA 961
```



```
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCAATTTTATCCAAAGTATATAAAACTCTCCCT 60
Db |||||
5940 ATGAAGACCATATGCTCTTTATCAGTCTTATCTTAATCCAAAGGATATAAAACTCTCCCT 5999
Qy 61 GTTTTAGAAATTTGCTAGTAAACAGCCCAACTCAAGATGTAGATTCAGTGTCTCCGGAAAC 120
Db |||||
6000 ATTTTAGAGATAGCTAGTAAACATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 6059
Qy 121 CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
Db |||||
6060 CTCAGAAGACAGAGATGTTCCATCTGATGGGATTTCAACATGAGTGGGCAAAAAGTTGCT 6119
Qy 181 GATTCCCTTTTGGAAAGCATCTAAAGCATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db |||||
6120 GATTCCCTTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 6179
Qy 241 GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 300
Db |||||
6180 GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 6239
Qy 301 TCTGGAATATCTTCTGCTCTGATGCTCCAGTAAATATCCGCGATTAACCTAAATGTAAA 360
Db |||||
6240 TCTGGAATATCTTCTGCTCTGATGCTCCAGTAAATATCCGCGATTAACCTAAATGTAAA 6299
Qy 361 ACTGTTTCATATATTCAGAGGTCAAAACCTCATGCAAGGGGATTTGCCCTCCATTTGTGG 420
Db |||||
6300 ACTATCCATCATATTCAGAGGTCAAAACCTCATGCGCAAGGGATGCGCCCTCCATTTGTGG 6359
Qy 421 GGGCATTTTCTTGATATGTCGGTTCCTCTCAACAATGTACCGAGGCAAGCTTTC 480
Db |||||
6360 GGAGCATTTTCTTGATATGTCGGTTCCTCTCAACAATGTACCGAGGCAAGCTTTC 6419
Qy 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATGAAGCAGTTCACAGAAATGATTTTTTCT 540
Db |||||
6420 ACTGAAGGAATATAGCAGCTATGATTTGTTAATGAAGCAGTTCACAGAAATGATTTTTTCT 6479
Qy 541 AGGCAAGGACAGGTATTCGTCAATGAACCTTGACCTCCACCAATATAATATTTGGACAGC 600
Db |||||
6480 AGGCAAGGACAGGTATTCGTCAATGAACCTTGACCTCCACCAATATAATATTTGGACAGC 6539
Qy 601 AGCAATGAACGACAGGAATGATACGGGATTTTGGCATCTCCCAAGAAATACACTCC 660
Db |||||
6540 AACAAATGAACGACAGGAATGATACGGGATTTTGGCATCTCCCAAGAAATACACTCC 6599
Qy 661 ACAACCAATCAAAATGACCTCCATCTCTTAAACCTCCATCCCTGCCCAACAGTAACTCCG 720
Db |||||
6600 ACGAAGATCAAAATGATGCTCCGTCCAAATACCTCACCACCTGCCCAACAGTAACTCCG 6659
Qy 721 AGCAATTCATCTACAAATPACTCAAAATTAATATCTGCTAAATCTGGAACCTATGAACCCAA 780
Db |||||
6660 GAGATCAAAACCAACAGCACCCCAACTGATGCCACCACTCAACACCAACAGACCCCAAC 6719
Qy 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGAGAACAGGGGCCCCACACA 840
Db |||||
6720 AATGATGATGAGGACCTTATGAATCAATCCGGTTCAGGGTCCGGAGAACAGGAAACCTTATA 6779
Qy 841 ACTCTTAATGATGACCTGAACAGAAACAAATCGTCAACAAATATTTGTCACCTCTTCACTA 900
Db |||||
6780 ACTTCAGATGCGGTCACTTAAGCAAGGGCTTTTCATCAACATGCCACCTCCCTCCACCA 6839
Qy 901 CATCAAGCACTCAACATGAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAACT 960
Db |||||
6840 CAACCAAGCACGCCACAGCAAGAGGAAACCAACAGACCACTTCCCAAGGTACTGTGACT 6899
Qy 961 GAGCAATATGAGGACCCCAACACACACAGCAAGCTCTCTCAACAAATACTAATACA 1020
Db |||||
6900 GAACCAACCAAAACCAACACACAGGCAACAGCGTCCATGCCCCCCCCCAACCACTGCA 6959
Qy 1021 ACTCCCACTTAACACTCTCAAGTACAACTCTCAGTACTCTCTTCCCTCCCAACCGCAAC 1080
Db |||||
6960 ATCTCTACTTAACACCTCTCAAGAACAACTTCAGCACCTCTCTGTATCACTACAAAC 7019
```

RESULT 12

AY358025

LOCUS

DEFINITION

AY358025 19151 bp RNA linear VRL 27-AUG-2003
Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin, complete

```
Qy 1081 ATCACCATTAATGATACACAACTGAACTAGCAGAAAGCGAAGCAAAACCAATGCTCAGTTG 1140
Db |||||
7020 ACCACCAATTAACGACACACAGAGCACAGCCACTGAAAATGAACCAACCCAGTGCCCTCG 7079
Qy 1141 AACCAAACTCTAGATCCAAACAGAAAATCCCAACACAGACAGCAAGACACCAACAGACAAC 1200
Db |||||
7080 AAAACAACCTGCTCTCCACAGGAAATCTTACCAACAGCAAGAGCACTTAACACACGAA 7139
Qy 1201 AACATCATATGACGACATCAGATATAACAAAGCAAAACACCCACAATTTCTTCGGAT 1260
Db |||||
7140 GGGCCCAACCAACAGGCAACCAATATGACAAATGGGCAATTTAAACAGTCCCTCCCCACC 7199
Qy 1261 TCTAGTCGGAACACCGCCCTCTATATATCTTTAGAAAGAAACCGATCGATCCTCTGGAG 1320
Db |||||
7200 CCCAACCGGACACACAACTCTTGTATATTTCAAGAAAGAAACGAAATATCCTCTGGAG 7259
Qy 1321 GAAGGACATGTTCCCTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGGACCA 1380
Db |||||
7260 GAAGGACATGTTCCCTTTCTGACGGGTTAATAATGCTCCAAATGATTTGATCCA 7319
Qy 1381 GTTCCAAATACAAAACCAATCTTTGATCAATCTCTAGTTCTGTGCTCGGCTGAGGAA 1440
Db |||||
7320 GTTCCAAATACAAAACCAATCTTTGATCAATCTCTAGTTCTGTGCTCGGCTGAGGAA 7379
Qy 1441 GATCAACATGCTCCCTCCCAATATCAGTTTAACTTTATCTTTTCTTAATAATAATGAG 1500
Db |||||
7380 GATCAACATGCTCCCTCCCAATATCAGTTTAACTTTATCTTTTCTTAATAATAATGAG 7439
Qy 1501 AACACTGCTACTCTCGAGAAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db |||||
7440 AACACTGCTACTCTCGAGAAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGGAGC 7499
Qy 1561 GTTCAGAGGATGACTGCGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAAT 1620
Db |||||
7500 GTTCAGAGGATGACTGCGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAAT 7559
Qy 1621 GAAGGACTTTACACTGCTGTTTTTAATAAATCAAAACATTTGGTCTGCAAGTTGAG 1680
Db |||||
7560 GAAGGACTTTATACACTGCTGTTTTTAATAAATCAAAACATTTGGTCTGCAAGTTGAG 7619
Qy 1681 CGTCTAGCCTCAATCAAACTGCCAAATCTTTGAACTCTTTATTTAGAGTCACAACTGAGGAA 1740
Db |||||
7620 CGTCTAGCCTCAATCAAACTGCCAAATCTTTGAACTCTTTATTTAGAGTCACAACCGAGGAA 7679
Qy 1741 AGAACATTTCTTTAATCAATAGACATGCTATTGACTTTTACTCACAAGATGGGGAGGA 1800
Db |||||
7680 AGGACATTTCTTTAATTAATAGACATGCTTACTTCTTCTCACAAGTGGGGAGGA 7739
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGCAATCGGATAGAAAGACTTGTCCAAAATATT 1860
Db |||||
7740 ACATGCAAAAGTCTTGGACCTGATTTGCAATCGGATAGAAAGACTTGTCCAGGAATATT 7799
Qy 1861 TCAGAGCAAAATTTGACCAAAATTTAAAGAGCAAAACCAAAAGAGGGGACTGTTGGGCTG 1920
Db |||||
7800 TCGGAAACAAATTTGACCAAAATTTAAAGAGCAAAACCAAAAGAGGGGACTGTTGGGCTG 7859
Qy 1921 GGTGTAATGCTTGGACATCCGACTGGGGTCTTCTTACTAATTTGGGCAATTTTGTCTACTA 1980
Db |||||
7860 GGTGTAATGCTTGGACATCCGACTGGGGTCTTCTTACTAATTTGGGCAATTTTGTCTACTA 7919
Qy 1981 TTATCCATAGTGTCTTGAATGCTCTATCTCTGATATTTGCTGATCTTTTACTAATATATC 2040
Db |||||
7920 TTATCCATAGTGTCTTGAATGCTCTATCTCTGATATTTGCTGATCTTTTACCATAATATC 7979
Qy 2041 GGATAA 2046
Db |||||
7980 GGTAA 7985
```


SVMGDNCITLTLSPFPIDAPNDVQENAEELNARVAVELAITTGYSGIFLKPEFTFVH
SGPIYFKKQYLVNGVQIPQSLKRWRCQPLSDSIFDDLOQSLASIGSFERSGSTRH
IFPWRWATASHFNLMAINLLMNHGIFPLIGENIDISCFKKPLTFSEKUALIITPOVLGG
LSLNPBKFLFYRNSDPLTGLGFLKNALFELGFLFYLIIAKKPGCLWADSDVFNPP
LGLNVPGRREITITLQTVRENITITQNRRIINSLFHIGDLEQKYCEWLLSNPVM
SRFAADI.FSRTPSKRLQVLGTRTLTLASRTISLTTEGMLMLKRLGNRNKRWK
FSYIDALDDLSBSLESLCTVDVANGFLRAYSWSDVLKGRLLIGATLPCLEQPKVM
INLESDREQNLSSESELINLLPYDCKELRLGGSHDTNLYVSCALDRKVQKHP
VRLAWTIGNRPFYIGRTEDKIGTPELRVNCPSAAUKBAEMVSRLLWNTQGTADRE
KLLIPLNRVNDYQIVNPLPHYSGNIVHRYNDQYQHSFWANRMSNSTRAIRIS
TNTLKYAGGQQAIDSNIIFONTINGLAVDIALSLAKSSASNTFRMLNKCCOT
RHVSEYLFDPKDLVDNLKMDNELVNDPLCGLIKGLRGRVSRSTFLSLNKSCT
GSYDFPTIAWLTGETVGSIFSDSSQSDTPISSGCTKTFVTHFLVVPVIESIYVAG
ANLVESLSRLIKSINKLSDTLPLISSTIRNLSHRSLRILLOSTFRHELVLITLAAHI
PLISLMLGGAGKSSDAVLFATIQNFHNFSCMLKKGQSSLPWLVIFPSEGOO
LKPILKILQRLSDLSLDPKQVCKTLDATCCSDSFVWYPSKSTRTHYASLNRYED
KANKVRSPTSHLNCFLFSLSTSSVSNQVTSNYSKIYVHPEDIPETINRTKLRY
GSTALQMDIKMPSSEONVENCPSKODIRPKDNOKITKHDORVEKESQOQSPSD
NMOTLAVMHSNPSQOTIKSIDVHEDFASRVILNSKINNPNLTDCTINTNLLTTPG
TEFLDTPLOSSRYSSTPRBSLLSREQASYLYVDCSNIPISISLDPGRNMSDONIQ
MLINAYKRDHLHCFDSNQFCRTFVSSMSHVIKGLPPPELRLKAICLAEGEGGARL
LLKWKETDHLFNTLATDSQOEABILSGRVIPLMYNIDRLSALLSRELRLLTNLLTQ
ITDITNPLWDSVLYQIPEDSDILTDMAETTKDREQLYKTIYVNIWTRSPNPKIS
IKVFLDYEGTLFLMKNALQYGOVLKPYSNKNSWYLCSSKRRIORLOIDFP
DOVGIFLICKAMSRQORAIPLYWLAHIEKNYPASLHEPFLTLGPSPLESSCHRYTIPF
SEGKALPHKQYVYRQSKOHLHSLMLDYENNSPLLDLRNHFICSLRKITKYNYNDILK
LNLVKAVEKGNKSVQIVETLPPNHSVCIVHVDHECFGCEKRLLLKLDLFRNTKIAEQ
KLLNRVIGYILFFPFLGSKSRSLRA"
complement (19133. .19151)

primer_bind
ORIGIN

Query Match 70.5%; Score 1443.4; DB 14; Length 19151;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTCTCATAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db ATGAGACTACATGCCTCTTATCAGTCTTATCTTATTAATCCAAAGGATATAAAACTCTCCCT 6019
Qy 61 GTTTTAGAAATGTCTAGTAAACAGCAACCTCAAGATGTAGATGTGCTCGGGAACC 120
Db ATTTTGGAGATAGCCAGTAAACGATCAACCCCAAAATGTGGATTCGGTATGCTCGGGAAC 6079
Qy 121 CTCGAAAGACAGAAAGATGTTCTATCTGATGGGATTTACATGATGGGCAAAAAGTTGCT 180
Db CTCGAGAAAACAGAAAGATGTTCCATCTGATGGGATTTACATGAGCGGGCAGAAAAGTTGCT 6139
Qy 181 GATTCCCTTTTGGAGCATCTAAACGATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db GATTCCCTTTTGGAGGCATCAAGCGATGGGCTTTCAGGACAGGTGTACCTCTCAAGAA 6199
Qy 241 GTTGAGTATACGGAAGGAGAGAGCCAAAACATGTTTCAATATAAGTGTAAACAGACCT 300
Db GTTGAGTACACGGAAGGGAGAGGCCAAAACATGCTACATATAAGTGTAAACGATCCC 6259
Qy 301 TCTGGAATAATCCTTGCTGCTGATCTCCCAAGTAAATFCCCGGATATACCTTAATGTAAA 360
Db TCTGGAATAATCCTTGCTGCTGATCTCCCAAGTAAATFCCCGGATATACCTTAATGTAAA 6319
Qy 361 ACTGTTCAATATATCAAGGTCAAAACCTCATGACAGAGGGATTCGCCCTCAATTTGTGG 420
Db ACTATCCATCAATATCAAGGTCAAAACCTCATGCGCAGGGGATTCGCCCTCAATTTGTGG 6379
Qy 421 GGGGCATTTTTCTTGATGATCGGCTTCCCTCTACAAACATGTATACCGGCAAGGTCTTC 480
Db GGAGCATTTTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6439
Qy 481 ACTGAAGAAATATAGCAGCTATGATGTTAATAGACAGTTCAAGAAATGATTTTTTCT 540
Db ACTGAAGGAAACATAGCAGCTATGATGTTAATAGACAGTTCAAGAAATGATTTTTTCTCG 6499
Qy 541 AGGCAGGACAGGTTATCGTACATGAACTTGACCTCCACCATAATATATTTGGACAGC 600
Db AGGCAGCTTTATACCGCAGGTTTAAATAAATAATCAAAACAATTTGGTCTCGCAGGTTGAGG 7639

Db 6500 AGGCAAGGACAAAGGTCACCATGAACTCTGACTTCTACTAATAAATATTGGACAAGT 6559
Qy 601 AGCAATGAAACGACAGAAATGATACGGGATGTTTTGGCATCTCCCAAGAAATACAATCC 660
Db 6560 AGCAACGGAACGCAACAAATGACATCGGATGTTTGGTACTCTTTCACGAGTACAATTTCT 6619
Qy 661 ACACCAATCAAAACATCGCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
Db 6620 ACGAAGAACCAAAACATGTCCTGCTTAAACACCCCCACCCGCCACAGCCGCTCCG 6679
Qy 721 AGCATTCACCTCTACAAATATCTCAAAATTAATCTCTGCTAAATCTGGAACCTATGAACCCAACT 780
Db 6680 GAATCAACCCACACAGCACCCCAACCGATGCGCACTAGACTCAACACCAACACCAAC 6739
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGGAAACAGGGGCCCCACACA 840
Db 6740 AGTGATGATGAGGATCTCAACAATCCGGCTCGGGGTCTGGGGAACAGGAACTCTATACG 6799
Qy 841 ACTCTTAATGTAGTCACTGAAACGAAACAACTCGTCAACAAATTTGTCACACTCTTCACTA 900
Db 6800 ACTTCTGATGGGTCACTTAAGCAAGGGCTTTTCATCAACAAATGCCACCCACTCCCTCACCG 6859
Qy 901 CATCCAAAGCACCTCAACAACATGACGAAAAACAGTACGAATCCTTCCCGACATGCTGTAAC 960
Db 6860 CAACGAGCAGCCACAGCAAGGAGGAAACACACAAACCACTCCCAAGACGCTACAAC 6919
Qy 961 GAGCAAAATGGAACCGACCCCAACAAACACACAGCAACAGCTCTCAACAATACTAATAACA 1020
Db 6920 GAACTTGATAACCAACACAACTGCAACACCGCGCCACCGCTCCCAACACACCAACACA 6979
Qy 1021 ACTCCCACTTATACACTCTCAAGTACAACTCAGTACAACTCTCTCCCTCCCAACCGCAAC 1080
Db 6980 ATCTCCACCAACAAACCTCCAAACACAACTCAGCACTCTCCGACCAACCAACAAAC 7039
Qy 1081 ATCAACCAATATGATACAAACGCTGAACTAGCAGAAACGCAACAAACCAATGCTCAGTTG 1140
Db 7040 ACCACCAATCCCAACACACAAAGCATGGCCACTGGAATGAGAAACCAAGTGCCTCCCG 7099
Qy 1141 AACCAAACTCTAGATCCAAAGAAAAATCCCAACACAGGACAGGACACCAACAGCAACACC 1200
Db 7100 AAAGCAACCTCGCTCCCAACAGAAAAATCTTACCAACAGAAAAAGAGCACCACAAATACAAA 7159
Qy 1201 AACATCATGACGACATCAGATATTAACAGCAAAACACCCCAACAAATTTCTTCTCCGAT 1260
Db 7160 AGCCCCACCAATGGAACCAATATCAACAAATGGAACATTTCTCACTAGTCTCTCTCCACC 7219
Qy 1261 TCTAGTCCGACCAACCGCTCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 7220 CCCAACTCGACTACAAACATCTTATATATTTAGGAGGAAACGAGATATCTCTGGAGG 7279
Qy 1321 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGACCCA 1380
Db 7280 GAAGGCGACATGTTCCCTTTCTAGATGGGTAAATAATGCTCCAAATGATTTTGATCCA 7339
Qy 1381 GTTCCAATACAAAAACAATCTTTGATGAATCTCTAGTCTGGTGCCTCGCGCTGAGGAA 1440
Db 7340 GTTCCAATACAAAAACAATCTTTGATGAATCTTTCTAGTCTGGTGCCTCAACTGAGGAA 7399
Qy 1441 GATCAACATGCTCTCCCAACATATAGTTTAACTTTTATCTTATTTTCTTAAATAATAATGAG 1500
Db 7400 GATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7459
Qy 1501 AACACTGCCTACTCTGGAGAAAAATGAGAAATGATGATGATGATGATGATGATGATGATG 1560
Db 7460 AACACTGCCTACTCTGGAGAAAAATGAGAAATGATGATGATGATGATGATGATGATGATG 7519
Qy 1561 GTTCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 7520 GTTCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7579
Qy 1621 GAAGGACTTTTACATGCTGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1680
Db 7580 GAAGGACTTTTATACCGCAGGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7639

```
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 7640 CGTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTACTAAGGGTCACAACCGAGGAA 7699
QY 1741 AGAACATTCTCTTAATCAATAGACATGCTATTGACTTCTTACTCAACAAGATGGGGAGGA 1800
Db 7700 AGAACATTCTCTTAATCAATAGACATGCTATTGACTTCTTACTCAACAAGATGGGGAGGA 7759
QY 1801 ACATGCAAGTGCCTGGACCTGATTGTGTCATCGGGATAGAGACTTCTCCAAAAATATT 1860
Db 7760 ACATGCAAGTGCCTGGACCTGATTGTGTCATCGGGATAGAGACTTCTCCAGAAAAATATT 7819
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 7820 TCAGAGCAAGATTGACCAAAATCAAGAGGACGAAACAAAAAGAGGGGACTGGTTGGGGCTG 7879
QY 1921 GGTGTAATGAGTGCATCCGACTCGGGGTGTTCTTACTAACTTGGGGCAATTTGCTACTA 1980
Db 7880 GGTGTAATGAGTGCATCCGACTCGGGGTGTTCTTACTAACTTGGGGCAATTTTACTACTA 7939
QY 1981 TTATCCATAGCTGCTTCAATGCTCTATCCTGATTTCTGCTATCTTACTAATATATC 2040
Db 7940 TTGTCATAGCTGCTTCAATGCTCTATCCTGATTTCTGCTATCTTACTAATATATC 7999
QY 2041 GGATA 2045
Db 8000 GGATA 8004

RESULT 13
AF005733
LOCUS
DEFINITION
Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin
glycoprotein precursor (GP) gene, complete cds.
ACCESSION
AF005733
VERSION
AF005733.1
KEYWORDS
GI:2459875
SOURCE
Lake Victoria marburgvirus
ORGANISM
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE
1 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
VARIATION
Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL
Virology 240 (1), 138-146 (1998)
MEDLINE
98110148
PUBMED
9448698
AUTHORS
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE
Direct Submission
Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
Location/Qualifiers
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strains="M/S.Africa/Johannesburg/1975/Ozolin"
/db_xref="taxon:11269"
1..2046
/gene="GP"
1..2046
/gene="GP"
/feature="viral surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40458.1"
/db_xref="GI:2459876"
EDVHLMGFTLSQKVADSPLEASKRWAFRTGVPPKNVYTBEGEAKTCYNISVTDPSG
```

```
KSLLDPPTNRYDPKCTIHHIQGNPHAGGIALHLWGAFFLVRIASTTMYRGKVF
TEGNIAAMIYNKTVHVKMIFSRQGGYRHMNLTSNKYNTNSNGTQINDTICGLTLEY
NSTKNOTCAPSKTPPPPTAPPEIKPTSTPTDRLNTNPNSDEDLTSGSGSGBEQ
ELVTSDAVTKQGLSSTMPPTPOGCTPOGGNNTNHSODATTELDONTTAAOPTPT
SHNTTISTNNTSKHNLSTLSAPPONTNNTOSMTVENKTSAPPKATLPTENPTT
EKSTNNTKSTTLPTNTINGHPTSPSTPNSTTQHLIYFRKRKRILILREGDMFPFLDG
LINAPIDFVPVNTKTIIDSSSSGASTEEDQHASNSILTSLSYLPHSTSGNTATYSGEN
ENDCDALRWSQVEDDLAAGLSWIFLPGIEGLYTAGLTKNNLVCRLRLANOT
AKSLELLRVTTEBRTPSLINRHAIDFLTRWGTCKVLGDCGICGEDLSEI
DQIKDKRQKSGTGNGLGKWTSDWGLVTNLGILLLSIAVLIALSCICRIPTKYIG"
1..81
/gene="GP"
82..2043
/gene="GP"
/product="glycoprotein"
1930..2019
/gene="GP"
/note="encodes transmembrane anchor peptide"

ORIGIN
Query Match 70.4%; Score 1441.2; DB 14; Length 2046;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 1668; Conservative 0; Mismatches 378; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAATCCAAAGTATATAAACTCCCT 60
Db 1 ATGAGGACTACATGCCTCTTTATCAGTCTTATCTTAATCCAAAGGATATAAACTCCCT 60
QY 61 GTTTTAGAATTGCTAGTAAACAGCCAACTCAAGATGTAGATTGATTCAGTGTCTCCGGAACC 120
Db 61 ATTTTGGAGATAGCCAGTAACGATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT 120
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAAACACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGGCGGACAGAAAGTTGCT 180
QY 181 GATTCCCTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTGGAGGATCCCAAGCATGGGCTTTTCAGGACAGGTGTTACCTCCTAAGAA 240
QY 241 GTTGATGATAGGAAAGGAGAGAGCCAAACATGTTACATATATAGTGTACAGACCTT 300
Db 241 GTTGATGATAGGAAAGGAGAGAGCCAAACATGTTACATATATAGTGTACAGACCTT 300
QY 301 TCTGAAAAATCTCTTGCTGCTGATCCCTCCAGTAAATATCCCGCATTAATCCCTAAATGTAAA 360
Db 301 TCTGAAAAATCTCTTGCTGCTGATCCCTCCAGTAAATATCCCGCATTAATCCCTAAATGTAAA 360
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCCCTCATGACAGGGGATTCGCTCCATTTGTGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGACAGGGGATTCGCTCCATTTGTGG 420
QY 421 GGGGCAATTTTCTTGATGATCGGTTGCTCTACAAATGTACCGAGGCAAGCTCTTC 480
Db 421 GGAGCAATTTTCTTGATGATCGGTTGCTCTACAAATGTACCGAGGCAAGCTCTTC 480
QY 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGCAGTTTCACAGAAATGATTTTTTCT 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGCAGTTTCACAGAAATGATTTTTTCT 540
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTTCACCTCCACCAATATAATATATGGCAAGC 600
Db 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTTCACCTCCACCAATATAATATATGGCAAGC 600
QY 601 AGCAATGAACGACAGAAATGATCGGATGTTTGGCATCTCCCAAGATATACACTCC 660
Db 601 AGCAATGAACGACAGAAATGATCGGATGTTTGGCATCTCCCAAGATATACACTCC 660
QY 661 ACAACCAATCAAAATGCTCCCTCCATCTCTTAAACCTCCATCCCTCCGCGGACAGTAACTCCG 720
Db 661 ACAACCAATCAAAATGCTCCCTCCATCTCTTAAACCTCCATCCCTCCGCGGACAGTAACTCCG 720
QY 721 AGCATTCCTCTACAAATATCTCAAAATTAATCTGCTAAATCTGGAACTATGAAACCAAGT 780
```

```
Db 721 GAATCAACCAACAGACCCCAACCGATGCCATAGACTCAACACCAACCAACCAAC 780
Qy 781 AGCGAGATGAGGACCTTATGATTTTCGGGCTCAGGATCTGGAGAACAGGGGCCCAACACA 840
Db 781 AGTGATGATGAGGATCTCAACAATCCGGCTCGGGCTCGGGGAACAGGAACTCTATAGC 840
Qy 841 ACTCTTAATGTAGTCACTGACAGAAACAATCGTCAACAAATATGTCCACTCCTTCACATA 900
Db 841 ACTTCTGATGGGTCACTAAGCAAGGGCTTTTCATCAACAATGCCACCCCACTCCCTCACCG 900
Qy 901 CATCAAGCACTCAACAATGAGCAAAACAGTAGTACGAATCTTCCCGACATGCTGTAACT 960
Db 901 CAACGAGCAGCCACAGCAGAGGAGAAACAACAACCAACCACTCCCAAGACGCTACAACT 960
Qy 961 GAGCACAATGGAACCGACCCCAACAACAACAGCAGCAACGCTCTCTCAACAATACTAATAACA 1020
Db 961 GAACTTTGATAACCAACAACCAACAGTGCACAACCGCCGCTCCCAACAACACCAACCAACA 1020
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGGCAAC 1080
Db 1021 ATCTCCCAACAACACCTCCCAACAACCAACCTCAGCACCTCTCCGCAACCAACCAACAAAC 1080
Qy 1081 ATCAACCAATATGATACACAACGCTGAATAGCAGAAAGCGCAACAACCAATGCTCAGTTG 1140
Db 1081 ACCACCAATCCCAACAACAAGCATGTGTCACTGAAATGAGAAACCAAGTGCCTCCCGCCG 1140
Qy 1141 AACCAACTCTAGATCCCAACAGAAAAATCCCAACAGCAGACAGACACCAACAGACACAACC 1200
Db 1141 AAGCAACCTGCCTCCCAACAGAAAAATCTTACCAGAAAGAGACCAACCAACATACAAA 1200
Qy 1201 AACATCATGAGACATCAGATATATAAGAGAAACACCCCAACAATTTCTTCGGAT 1260
Db 1201 AGCCCCACCAATGGAAACCAATACACAATGAGCAATTTCACTAGTCCCTCTCCACACC 1260
Qy 1261 TCTAGTCGCAACCGCCCTCTATATACCTTTAGAAAGAACATCGATCTCTGGAGG 1320
Db 1261 CCCAATCGACTACAAACATCTTATATATTTTCTAGGAGAAACGAAGTATCTCTGGAGG 1320
Qy 1321 GAAGCGCATGTTCCCTTTCTCGATGGGTAAATAATGCTCCAAATGATTTTGACCCA 1380
Db 1321 GAAGCGCATGTTCCCTTTCTAGATGGGTAAATAATGCTCCAAATGATTTTGATCCA 1380
Qy 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTTCTAGTCTGGTCTTCAACTGAGGAA 1440
Db 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTTCTAGTCTGGTCTTCAACTGAGGAA 1440
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTTCTTATTTTCTTAAATATAATGAG 1500
Db 1441 GATCAACATGCTCCCAATATAGTTTAACTTTTCTTATTTTCTTAAATATAATGAG 1500
Qy 1501 AACACTGCTACTCTGGAGAAAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db 1501 AACACTGCTACTCTGGAGAAAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Qy 1561 GTTCAGGAGATGACCTGGCCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTTCAGGAGATGACCTGGCCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAATC 1620
Qy 1621 GAAGGACTTTACATGCTGTTTTTAAATTAATAATCAAAACAATTTGGTCTGCAAGTTGAGG 1680
Db 1621 GAAGGACTTTTACCGCAGGGTTTTAATTAATAATCAAAACAATTTGGTCTGCAAGTTGAGG 1680
Qy 1681 CGTCTAGCAATCAAACTGCAAACTCTTGGAACTCTTATTTGAGAGTCAACAATCTGAGGAA 1740
Db 1681 CGTCTAGCAATCAAACTGCAAACTCTTGGAACTCTTATTAAGGTGTCACAAACGAGGAA 1740
Qy 1741 AGAACAATCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAACAAGATGGGGAGGA 1800
Db 1741 AGAACAATTTCTTAATCAATAGACATGCTATTGACTTTCTACTCAACAAGATGGGGAGGA 1800
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATGAGAGACTTGTCCAAAAATATT 1860
```

```
Db 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGAATAGAGACTTGTCCAGAAATATT 1860
Qy 1861 TCAGAGCAAAATTGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG 1920
Db 1861 TCAGAACAGATTTGACCAAAATCAAGAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG 1920
Qy 1921 GGTGGTAAATGGTGACATCCGACTGGGGTCTTCTTACTAACTTTGGGCAATTTTGGTACTA 1980
Db 1921 GGTGGTAAATGGTGACATCCGACTGGGGTCTTCTTACTAACTTTGGGCAATTTTACTACTA 1980
Qy 1981 TTATCCATAGCTGCTTGTGATTTGCTCTATCCTGTATTTGTCTATCTTTTACTAAATATATC 2040
Db 1981 TTGTCCATAGCTGCTTGTGATTTGCTCTATCCTGTATTTGTCTATCTTTTACTAAATATAT 2040
Qy 2041 GGATAA 2046
Db 2041 GGATAA 2046

RESULT 14
AX717717
LOCUS AX717717 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Musoke Glycoprotein 1 and Marburg virus strain Raven
Glycoprotein 2"

ORIGIN
Query Match 64.5%; Score 1319.6; DB 6; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATAATTTTCTGATTTAGTCTCAATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 1 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAATAATCTCCCC 60
Qy 61 GTTTTGAATTTGCTAGTACAGCCAACTCAAGATCTAGATTCAGTGTGCTCCGGAC 120
Db 61 ATTTTGAAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
Qy 121 CTCCAAAGACAGAAAGATGTTCTCATGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAAGACAGAAAGACGTCCTCATGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTGGGAAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTTCCTCCCAAGAAC 240
Db 181 GATTCCCTTTGGAGGCATCCAGCGATGGCTTTTCAGGACAGAGTGTACCTCCCAAGAAAT 240
Qy 241 GTTGAGTATACGAGAGGAGAGCCAAACATGTTACATATATAGTGAACAGACCCCT 300
Db 241 GTTGAGTATACAGAGGGGAGAGCCAAACATGTTACATATATAGTGAACAGGATCCC 300
Qy 301 TCTGGAATAATCTTGTGCTGGATCTTCCCAAGTAAATTCGCGGATTAACCTTAAATGTAAA 360
Db 301 TCTGGAATAATCTTGTGCTGGATCTTCTTACCAACATCCGCTGACTATCCGAATGCAAA 360
Qy 361 ACTGTTTCATATATTCAAGGTCAAAACCTTCATGACAGGGGATTTGCCCTCCATTTGTGG 420
```

361	Db		ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG	420
421	Qy	GGGGCATTTTTTCTTGATGATCGGGTGGCTCTACAACAATGTATACCGAGCGCAAGGCTCTTC	480	
421	Db	GGAGCATTTTTTCTTGATGATCGCATTTGGCTTCCACAACAATGTATACCGAGCGCAAGTCTTC	480	
481	Qy	ACTGAAGGAATATPAGCAGCTATGATTTGTTTAATAAGACAGTTTCACAGAATGATTTTTTCT	540	
481	Db	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTTCTCG	540	
541	Qy	AGGCAAGGACAAAGTTATCGTCACATGAACTTTGACCTCCACCAATAAATATTGACACAAGC	600	
541	Db	CGGCAAGGACAAAGGGTACCGGTCATATGAATCTGACTTCTACTATAAATAATTGACACAAGT	600	
601	Qy	AGCAATCAAAACGCAGAGAAATGATACGGGATGTTTTGCGATCTCTCCAAGAATAACAATCC	660	
601	Db	AGTAAACGGAAACGCAACGAATGACATGCGATGTTTTCGCGCTCTTCAAGAAATACAAATCT	660	
661	Qy	ACAAACAATCAAAACATGCGCTTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG	720	
661	Db	ACAAAGAACCAACATGTGCTCGGTCCAAATACCTCCACACTGCCACAGACCCGTCG	720	
721	Qy	AGCATTCACCTCTACAAATATCTCAAAATTAATCTGCTAAATCTGGAACTATGAAACCCAAGT	780	
721	Db	GAGATCAAACTCACAAAGCACCCCACTGATGCCACCAAACTCAATATCACGAGACCCAAGC	780	
781	Qy	AGCGACATGAGGACCTTATGATTTTCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840	
781	Db	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCCAACA	840	
841	Qy	ACTCTTAATGTAGTCACTGAAACAGAAACAATCGCTCAACAATTTGTCCACTTCCTTCACTA	900	
841	Db	ACTTCTGATGGGTTCACNAGCAAGGGCTTTCATCAACNAATGCCACCCACTCCCTCACC	900	
901	Qy	CATCCAAGCACCTCACAACTATGACGCAAAAACAGTACGAAATCTTTCCCGACATGCTGTAACT	960	
901	Db	CAACCAAGCACGCCACAGCAGGAGGAGAAACAACAACCACTTCCCAAGATGCTGTGACT	960	
961	Qy	GAGCACAAATGGAAACGGACCCCAACAACACACAGCAACGCTCTCTCAACATACTAATACA	1020	
961	Db	GACTAGACAAAATAACAACTTGCAACACGCTCCACGCTCCATGCCCCCTCATACACTACACA	1020	
1021	Qy	ACTCCCACCTATAACACTCTCAAGTCAAACTCAGTACTCTCTGCCCTCCAAACCCGCAAC	1080	
1021	Db	ATCTCTACTAACACACTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTACAAAC	1080	
1081	Qy	ATCACCAATATGATACAAACGTGAAGTACGAGAAAGCGCAACAAACCAATGCTCAGTTG	1140	
1081	Db	ACCACCAATGACAAACACACAGAGCACAACTCACTGAAAATGAGCAAAACCAAGTGCCTCTCG	1140	
1141	Qy	AACCAACTCTAGATCCAAACAGAAATCCCAACACAGACAGCAAGACACCAACAGCACACAC	1200	
1141	Db	ATAACAACTTCCTCCAAACGGGAAATCCCAACACAGCAAGACACCAAGAGACAGCAAAAA	1200	
1201	Qy	AACATCATATGACGACATCAGATATATCAAGAGCAAAACCCCAACAAATTTCTTCCTCGGAT	1260	
1201	Db	GGCCCCGCCACAAGGCACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCCACC	1260	
1261	Qy	TCTAGTCCGACAAACCGCCCTCCTATATATCTTTAGAAAGAAACGATCGATCCTCTGGAGG	1320	
1261	Db	CCGAGCTCGACTGCACAACACTCTTGATATTTTCAGNAGAAAGCGATCGATTTCTCTGGAAA	1320	
1321	Qy	GAAGGCGACATGTTCCCTTTTCTCGATGGGTTAATAANTGCTCCAAATGATTTTTGACCCA	1380	
1321	Db	GAAGGTGATATATTTCCCGTTTTTATGATGGGTTAATAATACTGAAATTTGATTTTGATCCA	1380	
1381	Qy	GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTCTGGTGCCTCGGCTGAGGAA	1440	
1381	Db	ATCCAAACACAGAAACAACTTTTATGATGAAATCTCCGACTTTAATACTTCAACTAATGAG	1440	
1441	Qy	GATCAACATGCCTCCCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATATAAATAGAG	1500	

Db	1441	GAACAACACACTCCCGGAATATCAGTTTAACTTTTCTCTTATTTCTCTGATAAAAATGGA	1501
Qy	1501	AACACTGCCTACTCTCGAGAAAAATGAGAATGATTGTGATGCAGAGTTTAAGAATTTGGAGC	1560
Db	1501	GATACTGCCTACTCTCGGGAACACGAGATGATTGTGATGCAGAGTTGAGGATTTGGAGT	1560
Qy	1561	GTTTCAGAGAGATGACCTCGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGATTT	1620
Db	1561	GTGCAGGAGACCATTTGGCGGCAGGGCTTAGCTGGATACCATTTTTTTGGCCCTTGGAACT	1620
Qy	1621	GAAGGACTTTTACACTGCTGTTTTTAATTTAAAAATCAAAACAAATTTGGTCTCGAGTTTCAGG	1680
Db	1621	GAAGGACTTTTACTGCGGTTTTTAATCAAAATCAGAACAATTTAGTTTGTAGTTGAGG	1680
Qy	1681	CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1740
Db	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTTAAGGGTCAACAACGAGGAA	1740
Qy	1741	AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAAGATGGGAGGA	1800
Db	1741	AGGACATTTTCTCTTAATCAATAGGCATGCAATTTGACTTTTGTCTACGAGTGGGCGGA	1800
Qy	1801	ACATGCABAAGTGTCTGACCTGATTGTCGATCGGATAGAAGACTTGTCCAAAAATATT	1860
Db	1801	ACATGCAAGGTGTAGGACCTGATTGTTGATAGGAATAGAGATCTATTCTAAAAATATC	1860
Qy	1861	TCAGAGCAATTCACCAAAATTAAGAAAGGACGAAACAAAAGAGGGGACTGTTGGGGTCTG	1920
Db	1861	TCAGACAAATTCACCAAAATCAGAAAGATGACAAAAGAGGAACTGGCTGGGTTCTA	1920
Qy	1921	GGTGGAATATGGTGGACATCCGACTGGGTGTTCTTACTTAACCTTGGGCATTTTGTCTACTA	1980
Db	1921	GGTGGCAATATGGTGGACATCTGACTGGGTGTTCTTACCACAAATTTGGGCATCTGCTACTA	1980
Qy	1981	TTATCCATAGCTCTCTTGATGCTCTATCTCTATCTGTATTTGTGCTATCTTTACTTAATATATC	2040
Db	1981	TTATCTATAGCTGTTCTGATTGCTCTGTCTGTATCTGTCTGATCTGTCTGATCTTCAATAATATC	2040
Qy	2041	GGATAA 2046	
Db	2041	GGATGA 2046	
RESULT 15			
LOCUS	AX717715	1841 bp	DNA linear PAT 15-APR-2003
DEFINITION	Sequence 3 from Patent WO02079239.		
ACCESSION	AX717715		
VERSION	AX717715.1	GI:29890726	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.		
TITLE	Chimeric filovirus glycoprotein		
JOURNAL	Patent: WO 02079239-A 3 10-OCT-2002;		
	U.S. Army Medical Research Institute of Infectious Diseases (US)		
FEATURES	Location/Qualifiers		
source	1. .1841		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="chimeric molecule between Marburg virus strain		
	Musoke Glycoprotein 1 and Ebola virus Zaire Mayinga strain		
	Glycoprotein 2"		
ORIGIN			
Query Match	38.1%;	Score 780;	DB 6; Length 1841;
Best Local Similarity	74.8%;	Pred. No. 7.4e-193;	
Matches 978;	Conservative	0; Mismatches 330;	Indels 0; Gaps 0;
Qy	1	ATGAAGACCATATATTTTCTGATGATGCTCATTTTAAATCCAAAGTATAAAAATCTCTCCCT	60

Query Match 38.1%; Score 780; DB 6; Length 1841;
Best Local Similarity 74.8%; Pred. No. 7.4e-193;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

ov 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAATCCAAAGTATAAAACTCTCCCT 60

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1071.87 Seconds
(without alignments)

11299.693 Million cell updates/sec

Title: US-10-066-506A-7

Perfect score: 2046

Sequence: 1 atgaagaccatattttct.....ttactaatatattcgataa 2046

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	10 ABT13453	Abt13453 DNA encod
2	1847.6	90.3	2046	10 ABT13456	Abt13456 DNA encod
3	1822	89.1	2247	4 AAD04043	Aad04043 Marburg v
4	1514.8	74.0	2103	8 ACC57764	Acc57764 Marburg v
5	1514.8	74.0	2164	4 AAD04042	Aad04042 Marburg v
6	1514.8	74.0	7778	10 ACC71550	Acc71550 VRC6701 (
7	1514.8	74.0	11460	3 AAZ87211	Aaz87211 VEE repli
8	1498.8	73.3	2100	3 AAZ51039	Aaz51039 Marburg v
9	1485.6	72.6	2051	10 ABT13454	Abt13454 DNA encod
10	1473.2	72.0	19112	13 ADM48345	Adm48345 Marburg v
11	1418.8	69.3	7005	10 ACC71551	Acc71551 VRC6702 (
12	1418.8	69.3	8256	10 ACC71552	Acc71552 VRC6710 (
13	1319.6	64.5	2046	10 ABT13452	Abt13452 DNA encod
14	834	40.8	6902	10 ACC71562	Acc71562 VRC6703 (p
15	780	38.1	1841	10 ABT13451	Abt13451 DNA encod
16	746.6	36.5	2252	10 ABT13450	Abt13450 DNA encod
17	149	7.3	18890	13 ADM48327	Adm48327 Reston eb
18	148.2	7.2	6885	10 ACC71538	Acc71538 VRC6302 (
19	148.2	7.2	6889	10 ACC71539	Acc71539 VRC6303 (
20	148.2	7.2	7001	2 AAX89795	Aax89795 DNA of pv

21	148.2	7.2	7002	10	ACC71536	Acc71536 VRC6300 (
22	148.2	7.2	7003	2	AAX59390	Aax59390 Plasmid p
23	148.2	7.2	7023	10	ACC71541	Acc71541 VRC6351 (
24	148.2	7.2	7036	10	ACC71537	Acc71537 VRC6301 (
25	148.2	7.2	8146	10	ACC71540	Acc71540 VRC6310 (
26	145.8	7.1	2961	10	ADF28971	Adf28971 Ebola vir
27	145.8	7.1	6913	10	ACC71531	Acc71531 VRC6101 (
28	145.8	7.1	8131	10	ACC71532	Acc71532 VRC6110 (
29	134.4	6.6	6940	10	ACC71535	Acc71535 VRC6202 (
30	134.4	6.6	7073	2	AAX59391	Aax59391 Plasmid p
31	134.4	6.6	7082	10	ACC71533	Acc71533 VRC6200 (
32	134.4	6.6	7087	10	ACC71534	Acc71534 VRC6201 (
33	134.4	6.6	8221	10	ACC71547	Acc71547 VRC6602 (
34	134.4	6.6	8338	10	ACC71546	Acc71546 VRC6601 (
35	134.4	6.6	10783	10	ACC71545	Acc71545 VRC6600 (
36	134	6.5	2039	10	ABT13455	Abt13455 DNA encod
37	134	6.5	2172	4	AAD04041	Aad04041 Ebola vir
38	134	6.5	2224	3	AAZ51038	Aaz51038 Ebola vir
39	134	6.5	2298	3	AAZ87189	Aaz87189 Ebola vir
40	134	6.5	2298	4	AAF76953	Aaf76953 Ebola vir
41	134	6.5	2298	12	ADL27460	Adl27460 Nucleotid
42	134	6.5	6467	10	ACC71530	Acc71530 VRC6052 (
43	134	6.5	6561	10	ACC71524	Acc71524 VRC6003 (
44	134	6.5	6624	10	ACC71523	Acc71523 VRC6002 (
45	134	6.5	6914	10	ACC71529	Acc71529 VRC6008 (

ALIGNMENTS

RESULT 1

ABT13453

ID ABT13453 standard; DNA; 2046 BP.

XX

AC ABT13453;

XX

DT 30-JAN-2003 (first entry)

XX

DE DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.

XX

KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

KW

immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX

OS Unidentified.

OS

Chimeric.

XX

FN WO200279239-A2.

XX

PD 10-OCT-2002.

XX

PF 31-JAN-2002; 2002WO-US003339.

XX

PR 31-JAN-2001; 2001US-0267522P.

XX

FA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX

PI Grogan CC, Hevey MC, Schmaljohn AL;

XX

DR WPI; 2003-040651/03.

XX

P-PSDB; ABJ18475.

DR

New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,

PT

useful for inducing an immune response against infection of different

PT

filoviruses, specifically against both Ebola and Marburg viruses.

XX

PS Claim 18; Page 78-80; 94pp; English.

XX

The invention relates to a chimeric filovirus glycoprotein (GP) protein

CC

comprising GP1 and GP2, where GP1 is from a filovirus different than that

CC

of GP2. The chimeric filovirus GP protein is useful for inducing an

CC

immune response against infection of different filoviruses, specifically

CC

against both Ebola and Marburg viruses by being used as a vaccine. This

CC

polynucleotide sequence represents a DNA encoding a chimeric filovirus

[illegible]

Db 2041 GGATAA 2046

RESULT 2
ABTI13456
ID ABTI13456 standard; DNA; 2046 BP.
XX
AC ABTI13456;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein RVN-GPI/GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GPI; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PP 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
XX P-PSDB; ABJ18478.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GPI and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Disclosure; Page 89-91; 94pp; English.
XX
CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GPI and GP2, where GPI is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
XX
SQ Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;

Query Match 90.3%; Score 1847.6; DB 10; Length 2046;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAATCTCCCT 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAATCTCCCT 60
Qy 61 GTTTTAGAATTTGCTAGTAACAGCCACCTCAAGATGTAGATTGCTGCTCCGGAACC 120
Db 61 GTTTTAGAATTTGCTAGTAACAGCCACCTCAAGATGTAGATTGCTGCTCCGGAACC 120
Qy 121 CTCAAAAGACAGAGATGTTTCATCTGATGGGATTTACATCTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAAAAGACAGAGATGTTTCATCTGATGGGATTTACATCTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTTGGAGCATCTAAACAGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTTGGAGCATCTAAACAGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Qy 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAAGACCCCT 300
Db 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAAGACCCCT 300

Qy 301 TCTGGAAAAATCCTTGCTGCTGGATCTCCAGTAATATCCGCGATTAACCTTAATGTAAA 360
Db 301 TCTGGAAAAATCCTTGCTGCTGGATCTCCAGTAATATCCGCGATTAACCTTAATGTAAA 360
Qy 361 ACTGTTTCATCATATTCAGGTCAAAAACCTCATGCAAGGGATTTGCCCTCATTTGTTGG 420
Db 361 ACTGTTTCATCATATTCAGGTCAAAAACCTCATGCAAGGGATTTGCCCTCATTTGTTGG 420
Qy 421 GGGGCATTTTCTTGATGATCGGTTGCTCTACAAACAATGTATCCGAGGCAAGGTCTTC 480
Db 421 GGGGCATTTTCTTGATGATCGGTTGCTCTACAAACAATGTATCCGAGGCAAGGTCTTC 480
Qy 481 ACTGAAGGAATATPAGCAGCTATGATTGTTAATAAGACAGTTCAAGATGATTTTCT 540
Db 481 ACTGAAGGAATATPAGCAGCTATGATTGTTAATAAGACAGTTCAAGATGATTTTCT 540
Qy 541 AGGCAAGGACAGGTTATCTCATGAACTTGACCTCCCAACCAATATATTTGACACAGC 600
Db 541 AGGCAAGGACAGGTTATCTCATGAACTTGACCTCCCAACCAATATATTTGACACAGC 600
Qy 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATACAACTCC 660
Db 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATACAACTCC 660
Qy 661 ACAAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
Db 661 ACAAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
Qy 721 AGCATTCACCTACAAATACCTCAAAATTAATCTGCTAAATCTGGAATATGAAACCCAGT 780
Db 721 AGCATTCACCTACAAATACCTCAAAATTAATCTGCTAAATCTGGAATATGAAACCCAGT 780
Qy 781 AGGACCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 781 AGGACCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAAACATCGTCAACAATATTTGCTCCACTCTTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAGAAAACATCGTCAACAATATTTGCTCCACTCTTCACTA 900
Qy 901 CATCCAGCCTCACAACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAACT 960
Db 901 CATCCAGCCTCACAACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAACT 960
Qy 961 GAGCACAATGGAACCGACCCCAACAAACACACAGCAACGCTCTCAACAATCTAATAACA 1020
Db 961 GAGCACAATGGAACCGACCCCAACAAACACACAGCAACGCTCTCAACAATCTAATAACA 1020
Qy 1021 ACTCCACCTTATACACTCTCAAGTACAACTCTCAGTACTCTCTCCCTCCAAACCCGCAAC 1080
Db 1021 ACTCCACCTTATACACTCTCAAGTACAACTCTCAGTACTCTCTCCCTCCAAACCCGCAAC 1080
Qy 1081 ATCAACCAATATGATACAAACGTTGAACTAGCAGAAAGCGAACAAACCAATGCTCAGTTG 1140
Db 1081 ATCAACCAATATGATACAAACGTTGAACTAGCAGAAAGCGAACAAACCAATGCTCAGTTG 1140
Qy 1141 AACCAACTCTAGATCCAAACAGAAAATCCCAACACAGCAAGACACCAACAGCACAACC 1200
Db 1141 AACCAACTCTAGATCCAAACAGAAAATCCCAACACAGCAAGACACCAACAGCACAACC 1200
Qy 1201 AACCATCTATGACGACATCAGATATATAAGCAAAACACCCCAACAATTTCTTCCGGAT 1260
Db 1201 AACCATCTATGACGACATCAGATATATAAGCAAAACACCCCAACAATTTCTTCCGGAT 1260
Qy 1261 TCTAGTCCGCAACCCGCCCTCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGAGG 1320
Db 1261 TCTAGTCCGCAACCCGCCCTCTCTATATATCTTTAGAAAGAAACGATCGATTTCTGAAA 1320
Qy 1321 GAGGCCACATGTTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTGTTGACCA 1380
Db 1321 GAGGTGATATATTTCCCGTTTTTAGATGGGTAAATAATGCTGAAATTTGATTTGATCCA 1380

```
QY 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTCTGTGTCCTCGGCTGAGAA 1440
Db ATCCCAACACAGAAACAAATCTTTGATGAATCCTCCAGCTTAAATACCTCAATGAG 1440
QY 1441 GATCAACATGCTCCGCCCAATATAGTTTAACTTATCTTATTTCTTAAATATAATGAG 1500
Db GAACAAACACATCCGCCCAATATAGTTTAACTTATCTTATTTCTTAAATATAATGAG 1500
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAATGATTGTGATGAGAGCTTAAAGAAATTTGGAGC 1560
Db GATACTGCTACTCTGGGGAACACGAGAAATGATTGTGATGAGAGCTTAAAGAAATTTGGAGT 1560
QY 1561 GTTCAGGAGATGACCTGGCCGACGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620
Db GTGCAGGAGGACGATTTGGCCGACGGCTTAGCTGATACCAATTTTTTTGGCCCTGGAATC 1620
QY 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
Db GAAGGACTTATACCTGCGGTTTTAATCAAAATCAGAACAAATTTAGTTTGTAGTTGAGG 1680
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGTCACAACTGAGGAA 1740
QY 1741 AGAACATTTCTTAAATCAATAGACATGCTATTGACTTTTCTACTCACAAGATGGGGAGGA 1800
Db AGGACATTTTCTTAAATCAATAGGATGCAATTGACTTTTGTCTTACGAGGTGGGGCGGA 1800
QY 1801 ACATCAAAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTCTCCAAAATATT 1860
Db ACATCAAAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTCTCCAAAATATT 1860
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGAGCTGTTGGGGCTG 1920
Db TCAGAGCAAAATGACCAAAATGACCAAAATGACCAAAAGGAGGAACTGGCTGGGGCTGA 1920
QY 1921 GGTGTAATGTTGACATCCGACTGGGGTGTTCTTACTAACTTGGGCATTTTGTCTACTA 1980
Db GGTGCAAAATGTTGACATCTGACTGGGGTGTTCTTACTAACTTGGGCATCTCTGCTACTA 1980
QY 1981 TTATCATAGTCTTGTGATGCTCTATCTGCTATCTGCTATTTCTGTAATCTTACTAATATATC 2040
Db TTATCATAGTCTTGTGATGCTCTATCTGCTATCTGCTATTTCTGTAATCTTACTAATATATC 2040
QY 2041 GGATAA 2046
Db GGATGA 2046

RESULT 3
ID AAD04043
AC AAD04043 standard; DNA; 2247 BP.
XX AAD04043;
XX AAD04043;
DT 02-JUL-2001 (first entry)
DE Marburg virus Ravn strain glycoprotein (GP) DNA.
XX Glycoprotein; GP; immune response; vaccine; antiviral;
KW type I transmembrane protein; ds.
XX Marburg virus.
XX Key Location/Qualifiers
FT CDS 97..2142
FT /*tag= a
FT /product= "Marburg virus Ravn strain glycoprotein (GP)"
XX
XX US6200959-B1.
XX
XX 13-MAR-2001.
XX
```

```
PF 04-DEC-1996; 96US-00760615.
XX
PR 04-DEC-1996; 96US-00760615.
XX
PA (POWD-) POWDERJECT VACCINES INC.
XX
PI Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX WPI; 2001-280564/29.
DR P-PSDB; AAE00708.
XX
PT Inducing an immune response to a Marburg or Ebola virus involves
PT delivering a genetic vaccine, which contains a genetic construct encoding
PT antigenic determinants for filovirus, using a particle acceleration
PT device.
XX
XX Claim 6; Col 33-40; 33pp; English.
PS
CC The patent discloses a method of inducing an immune response to Marburg
CC or Ebola virus glycoprotein which involves delivering a genetic vaccine,
CC containing a genetic construct encoding antigenic determinants for
CC filovirus, using a particle acceleration device. The genetic vaccine for
CC filovirus is created by joining a DNA sequence encoding at least a
CC portion of the filovirus glycoprotein to a promoter effective to promote
CC transcription of the DNA sequence. This method is useful for inducing
CC humoral, cell-mediated and secretory immune responses in the treated
CC individual. The present sequence is a DNA encoding Marburg virus Ravn
CC strain glycoprotein (GP), which is a type I transmembrane protein. This
CC sequence is used in the construction of genetic vaccine against
CC filoviruses
XX
SQ Sequence 2247 BP; 734 A; 518 C; 424 G; 571 T; 0 U; 0 Other;
Query Match 89.1%; Score 1822; DB 4; Length 2247;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1906; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Db 97 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 156
QY 61 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGATAGATTGATGTCCTCGGAACC 120
Db 157 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGATAGATTGATGTCCTCGGAACC 216
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTTACACTGAGTGGCAAAAAGTTGCT 180
Db 217 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTTACACTGAGTGGCAAAAAGTTGCT 276
QY 181 GATTCCCTTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 240
Db 277 GATTCCCTTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 336
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATATAGTTAACAGACCT 300
Db 337 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATATAGTTAACAGACCT 396
QY 301 TCTGAAAAATCCTTGTGCTGATCCTCCAGTAATATCCCGGATTTACCTTAAATGTAAA 360
Db 397 TCTGAAAAATCCTTGTGCTGATCCTCCAGTAATATCCCGGATTTACCTTAAATGTAAA 456
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAAACCTCATGCACAGGGGATTTGCCCTCCATTTGG 420
Db 457 ACTGTTTCATCATATTTCAAGGTCAAAAACCTCATGCACAGGGGATTTGCCCTCCATTTGG 516
QY 421 GGGGATTTTCTTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGCTCTTC 480
Db 517 GGGGATTTTCTTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGCTCTTC 576
QY 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGACAGTTTCACAGAAATGATTTTTTCT 540
Db 577 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGACAGTTTCACAGAAATGATTTTTTCT 636
```

QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGACCAAGC 600
DB 637 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGACCAAGC 696
QY 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC 660
DB 697 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC 756
QY 661 ACAACAAATCAAAACGTCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCCG 720
DB 757 ACAACAAATCAAAACGTCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCCG 816
QY 721 AGCATTTCACTTACAAATPACTCAAAATTAATCTGCTAAATCTGGAACCTATGAAACCCCAAGT 780
DB 817 AGCATTTCACTTACAAATPACTCAAAATTAATCTGCTAAATCTGGAACCTATGAAACCCCAAGT 876
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGAGGGCCCCACACA 840
DB 877 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGAGGGCCCCACACA 936
QY 841 ACTCTTAATGATGACATCGAACAGAAACAAATCGTCAACAAATATTTGTCACCTCTTCACTA 900
DB 937 ACTCTTAATGATGACATCGAACAGAAACAAATCGTCAACAAATATTTGTCACCTCTTCACTA 996
QY 901 CATCCAGCACCTCACACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAAC 960
DB 997 CATCCAGCACCTCACACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAAC 1056
QY 961 GAGCACAATGGAACCGACCCCAACAAACAGCAACGCTCTCCCAACAAATCTAATACA 1020
DB 1057 GAGCACAATGGAACCGACCCCAACAAACAGCAACGCTCTCCCAACAAATCTAATACA 1116
QY 1021 ACTCCACCTAATAACATCTCAAGTACAACTCTGATCTCTTCCCTCCCAACCGCAAC 1080
DB 1117 ACTCCACCTAATAACATCTCAAGTACAACTCTGATCTCTTCCCTCCCAACCGCAAC 1176
QY 1081 ATCCACCAATATGATACACACGCTGCACTAGCAGAAAGCGAACAACCAATGCTCAGTTG 1140
DB 1177 ATCCACCAATATGATACACACGCTGCACTAGCAGAAAGCGAACAACCAATGCTCAGTTG 1236
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGCAGACAGCAACCAACAGCAACAC 1200
DB 1237 AACCAACTCTAGATCCCAACAGAAATCCCAACAGCAGACAGCAACCAACAGCAACAC 1296
QY 1201 AACATCATGACGACATCATGATATACAGAGCAAAACCCCAACAAATTTCTCCGGAT 1260
DB 1297 AACATCATGACGACATCATGATATACAGAGCAAAACCCCAACAAATTTCTCTCCGGAT 1356
QY 1261 TCTAGTCCGACAAACCCGCTCTCTATATCTTTAGAAAGAACGATCGATCTCTGGAGG 1320
DB 1357 TCTAGTCCGACAAACCCGCTCTCTATATCTTTAGAAAGAACGATCTCTGGAGG 1416
QY 1321 GAAGCGACATGTTCCCTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 1380
DB 1417 GAAGTGATATATTTCCCTTTTATGATGGGTTAATAATGCTCCAAATGATTTTGATCCA 1476
QY 1381 GTTCCAAATACAAACAAATCTTTGATGAATCTCTAGTTCTGATGGCTCCGGCTGAGGAA 1440
DB 1477 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1536
QY 1441 GATCAACATGCTCCCAACAAATATGATTTAATCTTTATCTTTTCTTAAATATAAATGAG 1500
DB 1537 GAAACAACACACTCCCGGCAATATCAGTTTAACTTTCTTTCTTTCTGATGATGATGATGAGT 1596
QY 1501 AACCTGCTACTCTGAGGAAATCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1597 GATCTGCTACTCTGAGGAAATCAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1656
QY 1561 GTTTCAGGAGGATGACTCGCGCAGGCTCAGTTTGGATACCGTTTCTTTTGGCCCTCGAAT 1620
DB 1657 GTGAGGAGGAGGATGTTGCGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1716
QY 1621 GAAGGACTTTTACATGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGCGAGTTGAGG 1680

DB 1717 GAAGGACTCTTACTGCGGTTTAAATCAAAATCAGAACTTTAGTTTGTAGTTGAGG 1776
QY 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
DB 1777 CGCTTAGCTAATCAAACTGCTAAATCCTTGAGCTCTTGTTAAGGCTCACAAACGAGGAA 1836
QY 1741 AGACATCTCTTAAATCAATAGACATGCTATTGACTTTTCTACTCACAAGATGGGAGGA 1800
DB 1837 AGGACATTTTCTTAAATCAATAGGATGCAATGACTTTTGTCTTACGAGGTGGGGGGA 1896
QY 1801 ACATGCAAGTCTTGACCTGATTTGATCGGATAGAGACTTGTCCAAAAATATT 1860
DB 1897 ACATGCAAGTCTTGACCTGATTTGATCGGATAGAGACTTGTCTTAAANATATC 1956
QY 1861 TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAGAGGAGGACTGTTTGGGCTG 1920
DB 1957 TCAGACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGCTCTA 2016
QY 1921 GGTGGTAAATGGTGACATCGGACTGGGCTGTTCTTACTAACTTTGGGCAATTTGCTACTA 1980
DB 2017 GGTGGCAATGGTGACATCTGACTGGGCTGTTCTTCACTTGGGCACTCTCTCTACTA 2076
QY 1981 TTATCCATAGCTGCTTGTGATTTGCTCTATCTGATTTGCTGATCTTTTACTTAAATATATC 2040
DB 2077 TTATCTATAGCTGCTTGTGATTTGCTCTGCTGCTGATCTGCTGATCTTCACTTAAATATATC 2136
QY 2041 GGATFAA 2046
DB 2137 GGATGA 2142

RESULT 4
ACCS57764.
ID ACCS57764 standard; cDNA; 2103 BP.
XX
AC ACCS57764;
XX
DT 28-JUL-2003 (first entry)
XX
DE Marburg virus envelope glycoprotein coding sequence.
XX
KW Lenti-virus; pseudotyped virus; gene therapy; vector; gene; ss.
XX
OS Marburg virus.
XX
FH Key Location/Qualifiers
CDS 10..2055
FT /*tag= a
FT /product= "Envelope glycoprotein"
XX
FN WO2003035849-A2.
XX
PD 01-MAY-2003.
XX
PF 28-OCT-2002; 2002WO-US034545.
XX
PR 26-OCT-2001; 2001US-0353221P.
PR 26-OCT-2001; 2001US-0356436P.
XX
PA (MCCR/) MCCRAY P B.
PA (SAND/) SANDERS D A.
PA (DAVI/) DAVIDSON B L.
XX
PI Mccray PB, Sanders DA, Davidson BL;
XX
DR WPI; 2003-421416/39.
DR P-ESDB; ABR42242.
XX
PT New pseudotyped lenti-virus comprising a lentiviral capsid, a lipid bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful for eliciting an immune response against feline immunodeficiency virus infection.
PT

QY 1861 TCAGAGCAAAATTGACCAATTAATAAGGACGACAAACAAAGAGGGGACTGGTTGGGCTCG 1920
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1870 TCAGAGCAAAATTGACCAATTAATAAGGACGACAAACAAAGAGGGGACTGGTTGGGCTCG 1929
QY 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTTCTACTA 1980
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1930 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTTCTACTA 1989
QY 1981 TTATCCATAGCTGCTTGTGATGCTCTATCCCTGATATTTGTGCTATCTTTACTAATATATC 2040
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1990 TTATCCATAGCTGCTTGTGATGCTCTATCCCTGATATTTGTGCTATCTTTACTAATATATC 2049
QY 2041 GGATAA 2046
DB |||||||
2050 GGATAA 2055

RESULT 5
AAD04042
ID AAD04042 standard; DNA; 2164 BP.
XX
AC AAD04042;
DT 02-JUL-2001 (first entry)
XX
DE Marburg virus Musoke strain glycoprotein (GP) DNA.
XX
KW Glycoprotein; GP; immune response; vaccine; antiviral;
KW type I transmembrane protein; ds.
XX
OS Marburg virus.

Key Location/Qualifiers
FH 119. .2164
FT CDS /*tag= a
FT /*product= "Marburg virus Musoke strain glycoprotein (GP) "
XX

US6200959-B1.
XX
XX 13-MAR-2001.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX (POW-) POWDERJECT VACCINES INC.
XX
XX Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX
XX WPI; 2001-280564/29.
XX
XX P-PSDB; RAE00707.

Inducing an immune response to a Marburg or Ebola virus involves
PT delivering a genetic vaccine, which contains a genetic construct encoding
PT antigenic determinants for filovirus, using a particle acceleration
PT device.

Claim 5; Col 25-30; 33pp; English.
XX
XX The patent discloses a method of inducing an immune response to Marburg
XX or Ebola virus glycoprotein which involves delivering a genetic vaccine,
XX containing a genetic construct encoding antigenic determinants for
XX filovirus, using a particle acceleration device. The genetic vaccine for
XX filovirus is created by joining a DNA sequence encoding at least a
XX portion of the filovirus glycoprotein to a promoter effective to promote
XX transcription of the DNA sequence. This method is useful for inducing
XX humoral, cell-mediated and secretory immune responses in the treated
XX individual. The present sequence is a DNA encoding Marburg virus Musoke
XX strain glycoprotein (GP), which is a type I transmembrane protein. This
XX sequence is used in the construction of genetic vaccine against
XX filoviruses

SQ Sequence 2164 BP; 701 A; 536 C; 412 G; 515 T; 0 U; 0 Other;
Query Match 74.0%; Score 1514.8; DB 4; Length 2164;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTCTGATGTCTCATTTTAATCCTAAAGTATATAAACTCTCCCT 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
119 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAANAATCTCCCC 178
QY 61 GTTTTGAANAATGCTAGTAAACAGCCAACTCAAGATGTAGATTTCAGTGTCTCCGGAAAC 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
179 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 238
QY 121 CTCCTAAAGACAGAGATGTTTCATCTGATGGGATTTACATGATGGGCAAAAAGATTGCT 180
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
239 CTCCTAAAGACAGAGATGTTTCATCTGATGGGATTTACATGATGGGCAAAAAGATTGCT 298
QY 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAAGGACAGGTGTTCTCCCAAGAAC 240
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
299 GATTCCCTTTTGAAGCATCTAAAGGATGGCTTTTCAAGGACAGGTGTTCTCCCAAGAAC 358
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 300
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
359 GTTGAGTATACGAGAGGGGAGAGAGCCAAACATGCTACAATATAAGTGTAAACGATCC 418
QY 301 TCTGGAANAATCTTGTGCTGATGCTCTCCAGATTAATATCCGCGATTTACCTTAATGTAAA 360
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
419 TCTGGAANAATCTTGTGCTGATGCTCTTACCAACATCCGTGACTATCTTAAATGCAAA 478
QY 361 ACTGTTTCATCATATTCAGGTCAAAACCTTCATGACAGGGGATTCCTCATTTTGG 420
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
479 ACTATCCATCATATTCAGGTCAAAACCTTCATGACAGGGGATTCCTCATTTTGG 538
QY 421 GGGGCATTTTCTGTATGATCGGTTGCTCTTACAACATATGACGAGGCAAGTCTTC 480
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
539 GGAGCATTTTCTGTATGATCGCATTCCTCCACACATATGATACGAGGCAAGTCTTC 598
QY 481 ACTGAAGGAANAATAGCAGCTATGATTGTTTAATAGACAGTTTCAAGATGATTTTTTCT 540
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
599 ACTGAAGGGAACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTCTCG 658
QY 541 AGCGNAGGACAGGTTATCGTCAATGAACTTGACCTCCACCAATATAATTTGCAACAGC 600
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
659 CGGCAAGGACAGGTTATCGTCAATGAACTTGACCTCCACCAATATAATTTGCAACAGC 718
QY 601 AGCAATGAAACGACAGAAATGATCGGGATGTTTGGCATCTCTCCAAGATATACTTCC 660
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
719 AGTACGGAACGACAAACGATGACATCTGGATGTTTGGGGCTCTTCAAGATATACATCT 778
QY 661 ACAAAACATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
779 ACAAAACCAACATGCTGCTCCGTCCAAATATCTCCACCTGCCACAGCCCGTCCG 838
QY 721 AGCATTCATCTACAATACTCAAAATTAATCTGCTAAATCTGGAATATGAAACCCAGT 780
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
839 GAGATCAAACTCAACAGCACCCCACTGATGCTCCCAAACTCAATATCCACGACCCCAAGC 898
QY 781 AGCGACGATGAGGACCTTATGATTTTCGGCTCAGGATCTGGGAACAGGGGCCCCACACA 840
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
899 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGGAACGGAACCCACACA 958
QY 841 ACTCTTAATGTAGTCACTGAAACGAAACAATCGTCAACAATATTTGTCCACTCTTCACTA 900
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
959 ACTTCTGATGGGTCCACCAAGCAAGGGCTTTTCAACAATGTCACCACTCCCTCCACCA 1018
QY 901 CATCCAGGACCTCAACAATGACGAAACAGTACGATCTCTCCGACATCTGCTTAAC 960
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1019 CAACCAAGCAGCCACAGGAGGAGGAAACAACAAACCAATCTCCCAAGATGCTGTGACT 1078
QY 961 GAGCACAAATGGAACCGACCCCAACAAACACACAGCAAGCTCTCTCAACAACTAAATACA 1020
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1079 GAACCTAGACAAAATAACAACTGCAACACCGTCCATGCCCCCTCATACACTACCA 1138


```
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTCAGTACTCTTCCCTCCCAACCGCAAC 1080
Db 1139 ATCTCTACTAACACACCTCTCAAAACACAACTTCTGACACTCTCTCTGACCACTTACAAAC 1198
QY 1081 ATCACCATAATATACACACAGTGAATAGCAGAACGGAACAAACCAATGCTCAGTTG 1140
Db 1199 ACCACCAATGACACACACAGAGCAATCACTGAAATGAGCAACACAGTGCCCTCCG 1258
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGACAGCAAGACCAACAGACCAACC 1200
Db 1259 ATACACACCTGCTCCCAACCGGAATCCCAACAGCAAGAGACCAACAGCAGCAAA 1318
QY 1201 AACATCATCATGACGACATCAGATATAACAAGCAAAACCCCAACAAATTTCTTCGGAT 1260
Db 1319 GGCCTCCGACCAACAGCGCACCAACACAGCAAAATGAGCATTTTACCAGTCTCCCTCC 1378
QY 1261 TCTAGTCCGACAAACCGCGCTCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 1379 CCCAGCTCGACTGCAACAACATCTTGTATATTTTCAGAGAAAGCGAAGTATCTCTGGAGG 1438
QY 1321 GAAGCGCATGTTCCCTTTCTGATGGTTAATAAATGCTCCAAATGATTTTGACCA 1380
Db 1439 GAAGCGCATGTTCCCTTTCTGATGGTTAATAAATGCTCCAAATGATTTTGACCA 1498
QY 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 1440
Db 1499 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 1558
QY 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAAATAAATAG 1500
Db 1559 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAAATAAATAG 1618
QY 1501 AACCTGCTCTCTCGAGAAATCAGAAATGATGTCAGAGTTAAGAAATTTGGAGC 1560
Db 1619 AACCTGCTCTCTCGAGAAATGAGAAATGATGTCAGAGTTAAGAAATTTGGAGC 1678
QY 1561 GTTCAGGAGGATGACTCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 1620
Db 1679 GTTCAGGAGGATGACTCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 1738
QY 1621 GAAGGACTTTACATGCTGTTTTTAATAAATAAATAAATAAATTTGGCTGCGAGTTGAGG 1680
Db 1739 GAAGGACTTTACATGCTGTTTTTAATAAATAAATAAATAAATTTGGCTGCGAGTTGAGG 1798
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1740
Db 1799 GGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1858
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCAAGATGGGGAGGA 1800
Db 1859 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCAAGATGGGGAGGA 1918
QY 1801 ACATCAAAAGTCTGGAGCATGTTGTCATCGGGATAGAGACTTGTCCAAAATATT 1860
Db 1919 ACATCAAAAGTCTGGAGCATGTTGTCATCGGGATAGAGACTTGTCCAAAATATT 1978
QY 1861 TCAGAGCAAAATGACAAATTAAGAGGACGAAACAAAGAGGAGGACTGTTGGGCTG 1920
Db 1979 TCAGAGCAAAATGACAAATTAAGAGGACGAAACAAAGAGGAGGACTGTTGGGCTG 2038
QY 1921 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTAATGCGGCAATTTGCTACTA 1980
Db 2039 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTAATGCGGCAATTTGCTACTA 2098
QY 1981 TTATCCATAGTCTGCTGATGCTCTATCTGATATTTGTCGTATCTTTACTAATATATC 2040
Db 2099 TTATCCATAGTCTGCTGATGCTCTATCTGATATTTGTCGTATCTTTACTAATATATC 2158
QY 2041 GGATAA 2046
Db 2159 GGATAA 2164
```

```
RESULT 6
ACC71550
ID ACC71550 standard; DNA; 7778 BP.
XX
AC ACC71550;
XX
DT 10-JUL-2003 (first entry)
XX
DB VRC6701 (pVR1012-Marburg) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
PN WO2003028632-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030251.
XX
PR 01-OCT-2001; 2001US-0326476P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
PI WPI; 2003-371961/35.
XX
DR
XX
PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 185-187; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence has a Marburg glycoprotein (GP)
CC open reading frame, Musoke strain, cloned into a VRC6700 backbone
XX
SQ Sequence 7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;
Query Match 74.0%; Score 1514.8; DB 10; Length 7778;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60
Db 2034 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTATTAATCAAGGACAAAAAATCTCCC 2093
QY 61 GTTTTGAATTTGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTCTCCGGAACC 120
Db 2094 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGAATTCGTTATGCTCCGGAAC 2153
QY 121 CTCCAAAGACAGAGAGATGTTTCATCTGATGGGATTTACACTGAGTGGCAAAAAAGTTGCT 180
Db 2154 CTCGAAGACAGAGAGAGCTCCATCTGATGGATTCACACTGAGTGGCAAAAAAGTTGCT 2213
QY 181 GATTCCCTTTTGGAGGCAATCTAAACGATGGGCTTTCAGGACAGGTTGTTCTCTCCCAAGAAC 240
```


Db 2214 GATTTCCCTTTGGAGGCATCAAGCGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAA 2273
Qy 241 GTTGAGTATACGGAAGAGAGAGAGCCAAACATGTTACAAATAGAGTGAACAGACCC 300
Db 2274 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATAGAGTGAACAGATCCC 2333
Qy 301 TCTGGAATAATCCTTGCTGCTGGATCTCCCGAGTAATATCCCGGATTAACCCTAATATGTA 360
Db 2334 TCTGGAATAATCCTTGCTGCTGGATCTCCCGAGTAATATCCCGGATTAACCCTAATATGTA 2393
Qy 361 ACTGTTCAATCATATTCAGGTCAAAACCTCATGACAGAGGGATGGCCCTCAATTTG 420
Db 2394 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGAGGGATGGCCCTCAATTTG 2453
Qy 421 GGGGCATTTTCTTGATGATCGGGTGGCTCTCAACAAATGTACCGAGGCAAGGTCTTC 480
Db 2454 GGAGCATTTTCTTGATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGGTCTTC 2513
Qy 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAGACAGTTCAAGATGATTTTCT 540
Db 2514 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAGACAGTTCAAGATGATTTTCT 2573
Qy 541 AGGCAAGGACAGGTATTCGTCAATGACTGTGACCTCCACCAATATATGACAGC 600
Db 2574 CGGCAAGGACAGGTATTCGTCAATGACTGTGACCTCCACCAATATATGACAGC 2633
Qy 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCCAGGAATACAACTCC 660
Db 2634 AGTAACGGAAACGCAACGATGACACTGGATGTTTGGGCTCTTCAAGATACAACTCT 2693
Qy 661 ACAAAATCAAAATGATGCTCCATCTCTTAAACCTCCATCCCTGCGCACAGTAACCTCG 720
Db 2694 ACAAAATCAAAATGATGCTCCATCTCTTAAACCTCCATCCCTGCGCACAGTAACCTCG 2753
Qy 721 AGCATTCATCTACAAATCTCAATTAATCTGTAATCTGTAATCTGTAATCTGTAATCTG 780
Db 2754 GAGATCAAACTCAACAGCACCCCACTGATGCCCAAACTCAATACCAAGGACCCCAAGC 2813
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCAACACA 840
Db 2814 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACAGGACCCCAACACA 2873
Qy 841 ACTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 2874 ACTTCTGATGCGGTCAACAGCAGGGGCTTTCATCAACAAATGCCACCTCCCTCAACA 2933
Qy 901 CATCCAGCCTCACAACATGAGCAACAGTAGGATGATGATGATGATGATGATGATGATGATGAT 960
Db 2934 CAACCAAGCAGCCACAGCAAGGAGGAAACCAACCAATTCCTCAAGATGCTGTGACT 2993
Qy 961 GAGCACAATGGAACCGACCCCAACACACACAGCAACAGCTCTCAACAACTACTAATACA 1020
Db 2994 GAACAGCAAAATATACAACTGACACACAGCTGATGATGATGATGATGATGATGATGATGAT 3053
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 3054 ATCTCTACTAACAACACCTCAACACACAACTTCAGCACTCTCTGACCACTTACAAAC 3113
Qy 1081 ATCACCATAATATGATACACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 3114 ACCACCAATATGATACACAGCAGCAATCACTGAAATATGAGCAACCAAGTGGCCCTCG 3173
Qy 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGGACAGCAACCAACCAACAGCAAC 1200
Db 3174 ATAACACCTCTGCTCCCAACAGGAAATCCCAACAGGACAGCAACCAACAGCAAC 3233
Qy 1201 AACATCATATGACATCATGATATACAGCAACCAACCAACCAACCAACCAACCAACCAAC 1260
Db 3234 GGGCCCGCACAAACCGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3293
Qy 1261 TCTAGTCCGACAAACCGCCCTCTCTATATATCTTTAGAAAGAACGATCGATCTCTCGAGG 1320

Db 3294 CCAGCTCGACTGCACAAACATCTTGTATATTTTCAGAAAGAAAGCAAGTATCTCTCGAGG 3353
Qy 1321 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGAACCCA 1380
Db 3354 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGAACCCA 3413
Qy 1381 GTTCCAATATACAAACAAATCTTTGATGAATCCTCTAGTTCTGCTGCTCGGCTGAGGAA 1440
Db 3414 GTTCCAATATACAAACAAATCTTTGATGAATCCTCTAGTTCTGCTGCTCGGCTGAGGAA 3473
Qy 1441 GATCAACATGCTCCCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATAATAAG 1500
Db 3474 GATCAACATGCTCCCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATAATAAG 3533
Qy 1501 AACACTCCCTACTCTGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 3534 AACACTCCCTACTCTGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 3593
Qy 1561 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTGGCCCTGGAAAT 1620
Db 3594 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTGGCCCTGGAAAT 3653
Qy 1621 GAAGGACTTTTACACTGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGCAGGTTGAGG 1680
Db 3654 GAAGGACTTTTACACTGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGCAGGTTGAGG 3713
Qy 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTTGSAATCTTTATGAGAGTCACAACTGAGGAA 1740
Db 3714 CGTCTAGCCAAATCAAACTGCCAAATCCTTTGSAATCTTTATGAGAGTCACAACTGAGGAA 3773
Qy 1741 AGAATTTCTCTTAATCAATAGACATGCTATTTGATCTTCTACTCACAAGATGGGAGGA 1800
Db 3774 AGAATTTCTCTTAATCAATAGACATGCTATTTGATCTTCTACTCACAAGATGGGAGGA 3833
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATAAT 1860
Db 3834 ACATGCAAAAGTCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATAAT 3893
Qy 1861 TCAGAGCAAAATTCACAAATTTAAAAAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 1920
Db 3894 TCAGAGCAAAATTCACAAATTTAAAAAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 3953
Qy 1921 GGTGTAATAGGTGGACATCCGACTGGGTGTTCTTACTAATCTTGGGCAATTTGCTACTA 1980
Db 3954 GGTGTAATAGGTGGACATCCGACTGGGTGTTCTTACTAATCTTGGGCAATTTGCTACTA 4013
Qy 1981 TTATCCATAGCTGTTGATTTGCTCTATCCTGTAATTTGCTGCTATCTTTACTAAATATATC 2040
Db 4014 TTATCCATAGCTGTTGATTTGCTCTATCCTGTAATTTGCTGCTATCTTTACTAAATATATC 4073
Qy 2041 GGATAA 2046
Db 4074 GGATAA 4079

RESULT 7
AAZ87211
ID AAZ87211 standard; cDNA; 11460 BP.
XX
AC AAZ87211;
XX
DT 15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX
DE VEE replicon comprising Marburg virus genes.
KW VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV; filovirus; glycoprotein; GP; nucleoprotein; NP; structural protein; VP40; VP35; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.

OS Chimeric.
XX Key Location/Qualifiers
FH CDS 104..2182
FT /*tag= a
FT /product= "Marburg virus nucleoprotein (NP, AAY77128)"
FT 2944..3933
FT /*tag= b
FT /product= "Marburg virus structural protein VP35
FT (AAY77130)"
FT 4567..5478
FT /*tag= c
FT /product= "Marburg virus structural protein VP40
FT (AAY77129)"
FT 5940..7985
FT /*tag= d
FT /product= "Marburg virus glycoprotein (GP, AAY77127)"
FT 8864..9697
FT /*tag= e
FT /product= "Marburg virus structural protein VP30
FT (AAY77131)"
FT 10200..10961
FT /*tag= f
FT /product= "Marburg virus structural protein VP24
FT (AAY77132)"
XX WO200000616-A2.
XX
XX
PD 06-JAN-2000.
XX
XX
PF 21-JUN-1999; 99WO-US014174.
XX
XX
PR 29-JUN-1998; 98US-0091403P.
XX
XX (USNE-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
XX Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;
XX
XX WPI; 2000-160676/14.
DR P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132,
DR AAY77133.
XX
XX Novel Marburg virus vaccines used to induce an immune response against
FT the infection in nonhuman primates.
XX
XX
PS Claim 1; Page 56-57; 57pp; English.
XX
XX The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine
CC of the invention comprises a Venezuelan equine encephalitis virus (VEE)
CC replicon containing copies of the genes encoding Marburg virus (Musoke
CC strain) glycoprotein (GP), a deletion mutant GP (GP-delta-TM),
CC nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24.
CC The Marburg virus genes replace the VEE virus structural protein genes;
CC the result is a self-replicating RNA molecule that encodes its own
CC replicase and transcriptase functions, and in additional makes abundant
CC quantities of the Marburg virus proteins. When replicon RNA is
CC transfected into eukaryotic cells, along with two helper RNAs that
CC express the VEE virus structural proteins, the replicon RNA is packaged
CC into VEE virus-like particles by the VEE virus structural proteins, which
CC are provided in trans. Since the helper RNAs lack packaging signals
CC necessary for further propagation, the resulting VEE replicon particles
CC (VRPs) which are produced are infectious for one cycle but are defective
CC thereafter. On VNP infection of a cell, an abortive infection occurs
CC whereby the cell produces the Marburg virus proteins, is ultimately
CC killed by the infection, but does not produce any viral progeny. The VEE
CC replicon provides a potent tool for vaccination with Marburg virus
CC antigens. The replicons, vectors and constructs are used to produce
CC vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic
CC fever) in mammals, to elicit immune responses against Marburg antigens,
CC to confer protective immunity, and to reduce disease symptoms and reduce
CC the severity of disease. Studies of non-human primates vaccinated with
CC the replicon indicate that the vaccine will be efficient in protecting
CC humans against Marburg virus. Prior art Marburg virus vaccination

CC strategies have used formalin-inactivated Marburg virus, which are only
CC partially successful at protecting against Marburg virus infection. The
CC present sequence represents a cDNA corresponding to a VEE virus replicon
CC encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;
Query Match 74.0%; Score 1514.8; DB 3; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTATTCCAAAGTATAAAAACCTCTCCCT 60
Db 5940 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 5999
Qy 61 GTTTTGAAGAAATGCTAGTAAACAGCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC 120
Db 6000 ATTTTGAAGATAGCTAGTAAATATCAACCCCAAAATGTGGATTGCTATGCTCGGAAC 6059
Qy 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCGAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 6119
Qy 181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 6120 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 6179
Qy 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACAAATATAAGTGTACAGACCCCT 300
Db 6180 GTTGAGTATACGAGGGGGAGAGAGCCAAACATGTTACAAATATAAGTGTACAGACCCCT 6239
Qy 301 TCTGAAAAATCCTTGTGCTGGATCTCTCCAGTATATATCCGCGATTACCTCTAAATGTAAA 360
Db 6240 TCTGAAAAATCCTTGTGCTGGATCTCTCTACCAACATCCGTGACTATCTCTAAATGTAAA 6299
Qy 361 ACTGTTTCATATATCAAGGTCAAAACCTCATGTCACAGGGGGATGGCTCCATTTGTGG 420
Db 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGTCACAGGGGGATGGCTCCATTTGTGG 6359
Qy 421 GGGGCAATTTTCTGTATGATCGGTTGGCTCTACAACAATGTACCGAGGCAAGTCTTC 480
Db 6360 GGAGCATTTTCTGTATGATCGCATTTGGCTTCCACAACAATGTACCGAGGCAAGTCTTC 6419
Qy 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAAATGATTTTCT 540
Db 6420 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAAATGATTTTCT 6479
Qy 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATTAATTTGGACAAGC 600
Db 6480 CGGCAAGGACAGGTTATCGGTACATGAACTCCGTCTCTACTAATAAATATTTGGACAAGT 6539
Qy 601 AGCAATGAAAGGACAGAAATGATACGGGATGTTTGGCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCC 660
Db 6540 AGTAACGGAAGCAAAACGAAATGACATGATGTTTCCGGCTCTTCAAGAAATACAAATCT 6599
Qy 661 ACAACCAATCAAAACATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCCG 720
Db 6600 ACAAGAACCAAAACATGCTCCGTCCGTCCAAATACCTCCACCATGCCCCACAGCCGTCG 6659
Qy 721 AGCAATCACTCTACAATACTCAAAATTAATCTGCTAAATCTGGAACTATGAAACCAAGT 780
Db 6660 GAGATCAAACTCAACAGCACCCCAACTGATGCCAACCAAACTCAATACCAACGAGCCCAAGC 6719
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 6720 AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAACAGGAAACCCACACA 6779
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTCCACTCTCTTCACTA 900
Db 6780 ACTTCTGATGGGTCAACCAAGGAGGCTTTTCATCAACAATGCCACCCACTCTCTTCACTA 6839
Qy 901 CATCCAAGCACCTCACAAACATGAGCAAAAACAGTAGGAATCTCTCCGACATGCTGTAAC 960

Db 6840 CAACAGCAGCCACAGCAGGAGGAAACAAACAAACCAATCCCAAGATGCTGTGACT 6899
Qy 961 GAGCACAATGGAACCGACCCCAACACACAAACAGCAGCTCTCTCAACAACTACTAATACA 1020
Db 6900 GAACTAGACAAATAACACAACTGCAACACCGTCCATGCCCCCTCAACACTACTACCACA 6959
Qy 1021 ACTCCACCTATACACTCTCAAGTACACACCTCTAGTACTCTTCCCTCCACACCGCAAC 1080
Db 6960 ATCTCTACTAAACACCTCCAAACACAACTTACGACACTCTCTGACCACTTACAAAC 7019
Qy 1081 ATACCAATATATACACACAGTGAACTAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACACACACAGAGCACATCACTGAAATGAGCAACCAAGTGCCTCTCG 7079
Qy 1141 AACCAACTCTAGATCCAAAGAAAATCCCAACAGGACCAAGACACCAACAGCACCAACC 1200
Db 7080 ATAACAACCTGCTCCAAAGGGAATCCCAACAGCAGCAAGAGACACCAAGCAAAAAA 7139
Qy 1201 AACATCATGACGACATCAGATATACAGAGCAAAACACCCCAAAATTTCTTCCGGAT 1260
Db 7140 GGCCCGCCACACAGGCAACCAACACAGCAAAATGAGCAATTCACGAGTCTCTCCGCCACC 7199
Qy 1261 TCTAGTCGACAAACCGCCCTCTATATACCTTTAGAAAGAACGATCGATCCTCTGGAGG 1320
Db 7200 CCCAGCTGCACTGCACACACATCTGTATATATTCAGAAAGAGCAAGTATCTCTGGAGG 7259
Qy 1321 GAAGCGCAGCATGTTCCCTTTCTGGATGGGTAAATAATGCTCAAAATGATTTTGACCCA 1380
Db 7260 GAAGCGCAGCATGTTCCCTTTCTGGATGGGTAAATAATGCTCAAAATGATTTTGACCCA 7319
Qy 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTCTGTGCTCGGCTGAGGAA 1440
Db 7320 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTCTGTGCTCGGCTGAGGAA 7379
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTTATCTTTCTTAAATAAATGAG 1500
Db 7380 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTTATCTTTCTTAAATAAATGAG 7439
Qy 1501 AACACTGCTACTCTGGAGAAAATCAGAAATGATTTGATGACAGAGTTAAAGAAATTTGGAGC 1560
Db 7440 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGATGACAGAGTTAAAGAAATTTGGAGC 7499
Qy 1561 GTTCAGGAGGATGACTGGCCGAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAAT 1620
Db 7500 GTTCAGGAGGATGACTGGCCGAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAAT 7559
Qy 1621 GAAGGACTTTACATGCTGTTTAAATTAATAATCAAAACAAATTTGGTCTGAGGTTGAGG 1680
Db 7560 GAAGGACTTTACATGCTGTTTAAATTAATAATCAAAACAAATTTGGTCTGAGGTTGAGG 7619
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTTAGAGTCAACAACTGAGGA 1740
Db 7620 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTTAGAGTCAACAACTGAGGA 7679
Qy 1741 AGAATCTCTCTTAATCAATAGACATGCTATTTGATCTTTCTACTCAAGAATGGGGAGGA 1800
Db 7680 AGAATCTCTCTTAATCAATAGACATGCTATTTGATCTTTCTACTCAAGAATGGGGAGGA 7739
Qy 1801 ACATGCAAAAGTCTGGACCTGATTTGATCGGATAGAGACTTTGTCACAAAATATT 1860
Db 7740 ACATGCAAAAGTCTGGACCTGATTTGATCGGATAGAGACTTTGTCACAAAATATT 7799
Qy 1861 TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 1920
Db 7800 TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 7859
Qy 1921 GGTGTAATGAGTGGACATCCGACTGGGGGTGTTCTTAACTTTGGGCAATTTGCTACTA 1980
Db 7860 GGTGTAATGAGTGGACATCCGACTGGGGGTGTTCTTAACTTTGGGCAATTTGCTACTA 7919
Qy 1981 TTATCCATAGTCTTGATTTGCTATCTCTGATTTGCTGATCTTTTACTAAATATATC 2040

Db 7920 TTATCCATAGTCTTGATTTGCTATCTATCTCTGATTTGCTGATCTTTACTAAATATATC 7979
Qy 2041 GGATAA 2046
Db 7980 GGATAA 7985
RESULT 8.
AAZ51039
ID AAZ51039 standard; DNA; 2100 BP.
XX AAZ51039;
AC AAZ51039;
XX
DT 05-JUN-2000 (first entry)
XX
DB Marburg virus envelope glycoprotein DNA.
XX
KW Marburg virus envelope glycoprotein; viral glycoprotein;
KW pseudotyped retrovirus; MLV; Moloney murine leukaemia virus; antibody;
KW transduction; screening agent; immunological agent;
KW pharmacological agent; ss.
XX
OS Marburg virus.
XX
FH Key Location/Qualifiers
FT CDS 10..2052
FT /*tag= a
FT /product= "Marburg virus envelope glycoprotein"
XX
PN WO200008131-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017702.
XX
PR 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
XX
PA (PURD) PURDUE RES FOUND.
XX
PI Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
PI Fischbach MA;
XX
DR WPI; 2000-224030/19.
DR P-PSDB; AAY70075.
XX
PT Cells that produce inventive pseudotyped retroviruses having a broad host
PT range useful for introducing nucleotide sequences into target cells.
XX
PS Example 11; Page 60-61; 65pp; English.
XX
CC The present sequence encodes the Marburg virus envelope glycoprotein.
CC Plasmid pMBGP1 was produced by cloning into the plasmid pSP72 nucleotide
CC sequences corresponding to nucleotides 5931-8033 from the Marburg virus
CC genome. This is used for generation of stable cell lines transiently
CC producing Marburg-MLV (Moloney murine leukaemia virus) pseudotyped
CC retrovirus with a broad host range. These cells having different viral
CC glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
CC transduced with a desired ribonucleotide sequence. This is used to
CC identify screening agents effective in blocking viral entry into a cell.
CC These agents may be immunological agents like monoclonal or polyclonal
CC antibodies. The pharmacological agents include proteins, peptides or
CC various chemical agents. The pseudotyped retrovirus may be useful in
CC methods of identifying cell surface receptors that allow viral entry
XX
SQ Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;
Query Match 73.3%; Score 1498.8; DB 3; Length 2100;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1711; Conservative 0; Mismatches 332; Indels 3; Gaps 1;
Qy 1 ATGAGACCATATATTTTCTGATTTAGTCTCAATTTTAAATCCAAAGTATATAAACTTCCT 60

Db	10	ATGAAGACCA	CATGTTTCC	TTTTCCTTTAT	CAGTCTTAT	CTTAATCT	TAATTC	TAAGGGA	CAAAAAAT	TCTCCCC	69
Qy	61	GTTTTAGAAA	TGCTAGTAA	CAGGCAAC	CCTCAAGAT	GTAGATT	CAGTGTG	TCTCCGGAA	CC	120	
Db	70	ATTTAGAGAT	AGCTAGTAA	TATCAAC	CCCAAAAT	TGGGAT	CGGTAT	GCTCCGGAA	CT	129	
Qy	121	CTCCAAAGCA	GAGAAGAT	TGTTCA	TCTGANT	GGGATTTA	CACTGA	GTGGGCA	AAAAAGTTG	180	
Db	130	CTCCAGAAGA	CAGAAGAC	CGTCCAT	CTGATGGG	ATTCAC	TGATGGG	CACTCA	TGATGGGCA	AAAAAGTTG	189
Qy	181	GATTTCCCTT	TGGAGGAC	TCTA	AAACGAT	GGGCTT	TCAGG	CAGGTGT	TCTCCCAAG	240	
Db	190	GATTTCCCTT	TGGAGGAC	TCTA	AAACGAT	GGGCTT	TCAGG	CAGGTGT	TCTCCCAAG	249	
Qy	241	GTTGAGTAT	ACGGAAG	GAGAAAG	CCAAAC	CATGTTA	CAATAT	AAAGTGA	ACAGACCT	300	
Db	250	GTTGAGTAC	ACAGGGG	GAGAAAG	CCAAAC	CATGTTA	CAATAT	AAAGTGA	ACAGACCT	309	
Qy	301	TCTGGAAAA	TCTTGTCT	GTGGAT	TCCTCC	CAGTAA	TATFC	CGCGAAT	TACCTTAA	360	
Db	310	TCTGGAAAA	TCTTGTCT	GTGGAT	TCCTCC	CAGTAA	TATFC	CGCGAAT	TACCTTAA	369	
Qy	361	ACTGTTTCA	TATTC	AAAGTCA	AAACCT	CATGCA	CAGGGAT	TGCGCT	CCATTG	420	
Db	370	ACTATCCAT	CATATTC	AAAGTCA	AAACCT	CATGCA	CAGGGAT	TGCGCT	CCATTG	429	
Qy	421	GGGCACTTT	TTCTGT	ATGATCG	GGTTC	CTCTA	CAACAT	TGTAC	CGAGCAAG	480	
Db	430	GGAGCATTT	TTCTGT	ATGATCG	GGTTC	CTCTA	CAACAT	TGTAC	CGAGCAAG	489	
Qy	481	ACTGAAGGA	AAATAT	PAC	GAGCTAT	GATTTG	TTAATA	AGACAG	TTTCA	540	
Db	490	---GNAGG	GAACAT	PAC	GAGCTAT	GATTTG	TTAATA	AGACAG	TTTCA	546	
Qy	541	AGGCAAGCA	AGGTTAT	TC	TCA	ATGAA	CTTGA	CTCC	ACCAAT	600	
Db	547	CGGCAAGCA	AGGTTAT	TC	TCA	ATGAA	CTTGA	CTCC	ACCAAT	606	
Qy	601	AGCAATGAA	ACG	CAGAGAA	TGATAC	CGGATG	TTTGG	CACTC	TC	660	
Db	607	AGTACGGAA	ACG	CAGAGAA	TGATAC	CGGATG	TTTGG	CACTC	TC	666	
Qy	661	ACAAACAAT	CAAA	CATG	CCCTCC	ATCTCT	TAAAC	CTCC	AT	720	
Db	667	ACAAAGAAC	CAAA	CATG	CCCTCC	ATCTCT	TAAAC	CTCC	AT	726	
Qy	721	AGCATCTCT	CA	AAATAC	TAAAT	TA	CTG	TAA	CTG	780	
Db	727	GAGATCAAA	CTCTCA	CAAGC	ACCCAA	CTGATG	CCCA	CAAA	CTCA	786	
Qy	781	AGCGACGAT	GAGGAC	CTTAT	GATTTCC	GGCTC	CAGGAT	CTG	GAGAA	840	
Db	787	AGTGATGAT	GAGGAC	CTTAT	GATTTCC	GGCTC	CAGGAT	CTG	GAGAA	846	
Qy	841	ACTCTTAAT	GT	PAGTCA	CTG	TA	CAAGAA	CAAT	CTG	900	
Db	847	ACTTCTGAT	GGGTC	CA	CAAGAA	GGGCTT	TCAT	CA	CAAT	906	
Qy	901	CATCCAA	AGC	ACTTCA	CA	ATG	AGCA	AAA	CAGTAC	960	
Db	907	CAACCAAG	CA	CGCCAC	CA	AGCA	AGG	AGG	AA	966	
Qy	961	GAGCACA	AT	GGGAA	CGGAC	CCCA	CA	CA	CA	1020	
Db	967	GAACTAGCA	AAAAA	TAA	CACA	AACTG	CA	CA	CA	1026	
Qy	1021	ACTCCCA	CT	TA	AA	CA	CT	CT	CT	1080	
Db	1027	ATCTCTACT	AA	CA	CA	CT	CT	CT	CT	1086	
Qy	1081	ATCACCA	AT	AT	AT	AT	AT	AT	AT	1140	
Db	1087	ACCACCA	AT	AT	AT	AT	AT	AT	AT	1146	

QY	1141	AACACAACCTCTAGANTCAAAGAAAATCCCAACACAGGAAACAGACACCACACAGCACAAACC	1206
DB	1147	ATAACAACCCCTGCCCTCCAACGGGAAAATCCCACACACAGCAGAAAGAGACCCAGCAGCAAAAAA	1206
QY	1201	AACATCATCATGACGACATCAGATATAAACAAGCAAAACACCACCAAAATCTTCTCGSGAT	1266
DB	1207	GCCCCGCCAACAGCGGCACCAACACGACAATATGAGCATTTCAACAGTCCTCCCCCACC	1266
QY	1261	TCTAGTCCGACAACCCGCCCTCCTATATCTTTAGAAAGAAGACGTATCGATCCTCTGGAGG	1320
DB	1267	CCGAGCTCGACTGCACAACATCTTGTAATATTCAGAAAGAAAGCAAGTAGTATCCTCTGGAGG	1326
QY	1321	GAAGGACACATGTTCCCTTTTTCTGGATGGGTAAATAAAGCTCAATTGATTTTGACCCA	1380
DB	1327	GAAGGACACATGTTCCCTTTTTCTGGATGGGTAAATAAAGCTCAATTGATTTTGACCCA	1386
QY	1381	GTTCCTCAATACAAAAACAATCTTTTGATGATCTCTAGTTCTGGTGCTCGGCTGAGGAA	1440
DB	1387	GTTCCTCAATACAAAAACAATCTTTTGATGATCTCTAGTTCTGGTGCTCGGCTGAGGAA	1446
QY	1441	GATCAACATGCTCTCCCCCAATATTTAGTTTAACTTTATCTTTATTTCTTAATAATAAATGAG	1500
DB	1447	GATCAACATGCTCTCCCCCAATATTTAGTTTAACTTTATCTTTATTTCTTAATAATAAATGAG	1506
QY	1501	AACACTGCTACTCTGAGAAAAATGAGATGATTTGTGATCAGAGTTAAGAAATTTGGAGC	1560
DB	1507	AACACTGCTACTCTGAGAAAAATGAGATGATTTGTGATCAGAGTTAAGAAATTTGGAGC	1566
QY	1561	GTTCAGAGAGATGACCTGGCCGAGGCTCAGTTGGATACCGTTTTTTTGSCCTCGGAAT	1620
DB	1567	GTTCAGAGAGATGACCTGGCCGAGGCTCAGTTGGATACCGTTTTTTTGSCCTCGGAAT	1626
QY	1621	GAAGGACTTTACACTGCTGTTTTAAATPAAAAATCAAAACAATTTGGTCTCGAGGTTGAGG	1680
DB	1627	GAAGGACTTTACACTGCTGTTTTAAATPAAAAATCAAAACAATTTGGTCTCGAGGTTGAGG	1686
QY	1681	CGTCTAGCAATCAAACTGCCAAATCTTGGAACCTTTATTTGAGAGTCACAACTGAGGAA	1740
DB	1687	CGTCTAGCAATCAAACTGCCAAATCTTGGAACCTTTATTTGAGAGTCACAACTGAGGAA	1746
QY	1741	AGACATCTCCTTAATCAATAGACATGCTATTGACTTTCTCTCAACAGATGGGAGGA	1800
DB	1747	AGACATCTCCTTAATCAATAGACATGCTATTGACTTTCTCTCAACAGATGGGAGGA	1806
QY	1801	ACATGCAAAAGTGTCTGGACCTGATTTGTTCATCGGATAGAAGACTTGTCCAAAAATATT	1860
DB	1807	ACATGCAAAAGTGTCTGGACCTGATTTGTTCATCGGATAGAAGACTTGTCCAAAAATATT	1866
QY	1861	TCAGAGCAAAATGACCAAAATPAAAAAGGACGAACAAAAAGAGGGAAGTGGTGGGGTCTG	1920
DB	1867	TCAGAGCAAAATGACCAAAATPAAAAAGGACGAACAAAAAGAGGGAAGTGGTGGGGTCTG	1926
QY	1921	GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGCACTTTTGTCTACTA	1980
DB	1927	GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGCACTTTTGTCTACTA	1986
QY	1981	TTATCCATAGTGTCTTTGATTGCTCTATCTCTGATTTTGTGCTATCTTTACTAAATATATC	2040
DB	1987	TTATCCATAGTGTCTTTGATTGCTCTATCTCTGATTTTGTGCTATCTTTACTAAATATATC	2046
QY	2041	GGATAA	2046
DB	2047	GGATAA	2052

RESULT 9
ABT13454
ID ABT13454 standard; DNA; 2051 BP.
XX
AC ABT13454;
XX
DT 30-JAN-2003 (first entry)

XX DNA encoding a chimeric filovirus protein MBGV-GP1/GP2.
XX
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
XX Unidentified.
OS Chimeric.
XX
XX W0200279239-A2.
FN
XX
XX 10-OCT-2002.
PD
XX
XX 31-JAN-2002; 2002WO-US003339.
PP
XX
XX 31-JAN-2001; 2001US-0267522P.
PR
XX
XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
PA
XX
XX Grogan CC, Hevey MC, Schmaljohn AL;
PI
XX WPI; 2003-040651/03.
DR
XX P-PSDB; ABJ18476.
DR
XX
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
PT
XX
XX Disclosure; Page 81-84; 94pp; English.
PS
XX
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
CC
XX
XX Sequence 2051 BP; 655 A; 520 C; 400 G; 476 T; 0 U; 0 Other;
SQ

Query Match 72.6%; Score 1485.6; DB 10; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;

Qy 1 ATGAAGACCATATTTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAATCTCCCT 60
Db 10 ATGAAGACCATATTTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAATCTCCCT 69
Qy 61 GTTTTAGAATGCTAGTAACAGCAACCTCAAGATAGATTAGTGTGCTCCGGAAC 120
Db 70 ATTTTAGAGATAGTGTAGTAATTAATCAACCCCAAAATGTGGATGCTCCGGAAC 129
Qy 121 CTCCAAAGACAGAAGATGTTTCATGTGGGATTTACATGAGTGGGCAAAAATGTTGCT 180
Db 130 CTCGAGAGACAGAGAGCTCCTATCTGATGGGATTTACATGAGTGGGCAAAAATGTTGCT 189
Qy 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGAGTGTTCCTCCCAAGAAC 240
Db 190 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGAGTGTTCCTCCCAAGAAC 249
Qy 241 GTTAGATAG 300
Db 250 GTTAGATAG 309
Qy 301 TCTGAAATCCTTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 310 TCTGAAATCCTTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 369
Qy 361 ACTGTTTCATATATTCAGGTCAAAACCTCATGACAGAGGGGATGTCCTCCATTTGTGG 420
Db 370 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGAGGGGATGTCCTCCATTTATGG 429
Qy 421 GGGGATTTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

Db 430 GGAGCATTTTCTGTATGATCGCATTT-----ATGTACCGGCAAGATCTTCT 477
Qy 481 ACTGAAGAAATATAGCAGCTATGATTGTTTAATAAGACAGTTTCAAGAAATGATTTTCT 540
Db 478 ACTGAAGGAAACATAGCAGCTATGATTGTTCAATAAGACAGTGCACAAAATGATTTCTCG 537
Qy 541 AGGCAAGGACAAAGTTATGTCATGAACTTGTACCTCCACCAATAAATATTGACAGAC 600
Db 538 CGGCAAGGACAAAGGTTACCGTCAATGAAATCTGATTTCTATATAAATATTGACAACT 597
Qy 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTGTGSCATCTCTCCAAAGAAATACA 660
Db 598 AGTAAACGAAACGCAACGAATGACATCTGGATGTTTCGGCGCTCTTCAAGAAATACAT 657
Qy 661 ACAAAACATCAAAACATGCCCTTCCATCTCTTAAACCTCCATCTCCGCCCACAGTAAC 720
Db 658 ACAAAAGAACAAACATGTGCTCGTCCAAATACCTCCACCATGCCACACGCGCTCG 717
Qy 721 AGCATTCATCTACAAATACTCAAAATTAATCTGCTAAATCTGGAATCTGAACCCCAAGT 780
Db 718 GAGATCAAACTCAACAGCACCCCACTGATGCCACCAAACTCAATACCCAGGACCAAGC 777
Qy 781 AGGACATGAGGACCTTATGATTTTCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 778 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACAGGAACCCACACA 837
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTGTCCACTCTCTTCA 900
Db 838 ACTTCTGATCGGTCCACCAAGCAAGGGCTTTTCATCAACAATGCCACCTCCCTCACA 897
Qy 901 CATCAAGACCTCACAACATGAGCAAAACAGTAGAGAAATCTTCCGACATGCTGTAACT 960
Db 898 CAACCAAGCAGCCACACAGAGGAGGAAACAACAACCAATTTCCCAAGATGCTGTGACT 957
Qy 961 GAGCACATGGAACCGACCCCAACAACAACACAGCAAGCTCTCAACAATACTAATAACA 1020
Db 958 GAATAGACAAAATAACAACTGCAACACCTGCAACACCTGTCATGCCCTCTATAACACT 1017
Qy 1021 ACTCCCACTTAACTCTCAAGTACAACTCAGTACTCTCTTCCCTCCAAACCCGCAAC 1080
Db 1018 ATCTCTACTAACACACTCTCCAAACACACTTTCAGCACTCTCTCTGACCATTAACAAC 1077
Qy 1081 ATCAACCAATATGATACAAACGTTGAACTAGCAGAAAGCGGAAACAACAATGCTCAGTT 1140
Db 1078 ACCACCAATGACACACACAGAGCACAATCACTGAAATGAGCAAAACGAGTGCCTCG 1137
Qy 1141 AACCAACTCTAGATCCAAACAGAAATCCCAACACAGCAAGCAACCAACAGCACACACC 1200
Db 1138 ATAAACAACCTGCTCCAAACGGAATTCACCAACAGCAAGAGAGAGAGAGAGAGAGAG 1197
Qy 1201 AACATCATGACGACATCAGATATTAACAGCAAAACACCCCAACAAATTTCTTCCGGAT 1260
Db 1198 GGCCCCGCAACAGGCAACCAACAGCAAAATGAGCATTTTCAACAGTCTCTCCGCCAC 1257
Qy 1261 TCTAGTCCGCAACACCCGCTCTTATATATCTTTAGAAAGAAACGATCGATCTCTGAGG 1320
Db 1258 CCCAGCTCGATGCAACATCTTGTATATTTTCAAGAAAGAGCGATCGATCTCTGAGG 1317
Qy 1321 GAAGGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCA 1380
Db 1318 GAAGGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCA 1377
Qy 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCCTCGGCTGAGGAA 1440
Db 1378 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCCTCGGCTGAGGAA 1437
Qy 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATATAAATAG 1500
Db 1438 GATCAACATGCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATATAAATAG 1497
Qy 1501 AAGCTGCTCTCTGAGAGAAATGAGATGATGATGATGATGATGATGATGATGATGATG 1560

Qy	361	ACTGTTCAATCATATTCAAGGTCAAAACCCCTCATGCACAGGGGATGCGCCCTCCATTGTGTG	420
Db	6300	ACTATCCATCATATTCAAGGTCAAAACCCCTCATGCGCAAGGGATGCGCCCTCCATTGTGTG	6359
Qy	421	GGGCACTTTTCTGTATGATCGGTTGCTCTTCAACAAATGTACCGAGCAAGGTCCTTC	480
Db	6360	GGAGCATTTTTCCTGTATGATCGGATTTGCTCCACAACAATGTACCGAGGACAGATCTTC	6419
Qy	481	ACTGAAGAAATATAGCAGCTATGATTGTTAAATAAGACAGTTCACAGAATGATTTTCT	540
Db	6420	ACTGNAAGGACATAGCAGCTATGATTGTCAATAAGACAGTTCACAGAATGATTTTCTCG	6479
Qy	541	AGGCAAGGACAAGGTTATCGTCACATGAACTTGTGACCTCCACCAATAAATATCGACAAGC	600
Db	6480	AGGCAAGGACAAGGGGTACCGTCAATGAATCTGACTTCTACTAATAAATATCGACAAGT	6539
Qy	601	AGCAATGAAAGCGACAGAAATGATACGGGATGTTTGGCATCTCCAGAAATCAACTCC	660
Db	6540	AACAATGGAACACAAACGAATGACACTGGAATGCTTGGTGCTCTTCAAGAATACAACTCC	6599
Qy	661	ACAAACAATCAAAATCGCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCCG	720
Db	6600	ACGAAGAATCAAAATGCTGCTCCGTCCAAATAACCTCACACTGCCCCACAGCCGCTCCA	6659
Qy	721	AGCAATTCACCTCTACAAATACTCAAATTAATACTGCTAAATCTGGAACTATGAACCCAAAGT	780
Db	6660	GAGATCAAAACCCACACAGCACCCCAACTGATGCCACACACACTCAACACCCACAGACCCAAAC	6719
Qy	781	AGCAGCAGTAGAGACCTTATGATTTCCGGCTCAGGATCTGAGAGAAAGGGGGCCCCACACA	840
Db	6720	AATGATGATGAGGACCTCATAAATCCGGTTTCAGGGTCCGGAGAAACAGGAACCTTATACA	6779
Qy	841	ACTCTTAATGTAGTCACGTGACACAGAAACAATCGTCAACAATATTTGTCACCTCTTCACATA	900
Db	6780	ACTTCAGATGGGTGACTTAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA	6839
Qy	901	CATCCAAGCACCTCAACAACATGAGCAAAAACAGTACGAATCTCTCCCGACATGCTGTAACCT	960
Db	6840	CAACCAAGCAGCCACACGAAAGAGGAAACAAACACAGACCATTTCCCAAGGTACTGTGACT	6899
Qy	961	GAGCACAATGGAAACCGACCCCAACAACAACACAGCAACGGTCTCTCAACAATATTAATACA	1020
Db	6900	GAACCCAAACAAAACCAACACAAACGGCACAAACCGTCCATGCGCCCCCCCCACACACCACTGCA	6959
Qy	1021	ACTCCCACCTATACACTCTCAAGTACAACTCAGTACTCTTTCCTCCCTCCAAACCGCAAC	1080
Db	6960	ATCTCTACTAAACAACCTCCAAAGAACACTTTCAGCACCTCTCTGTATCACTACAAAC	7019
Qy	1081	ATCACCAATATGATACACAACGTGAACTAGCAGAAAAGCGAAACAAACCAATGCTCAGTTG	1140
Db	7020	ACCAACCAATTAGCACACACAGACACAGCCACTGAAATGAAACAAACAGTGCCTCCCTCG	7079
Qy	1141	AACACAACCTTAGATCCAAACAGAAAATCCCAACACAGGACAAGACACCAACAGCAACAAC	1200
Db	7080	AAAACAACCTTCCTCCAAACAGGAAATCTTTACACACAGCAAAAGACACTAACAACACGAAA	7139
Qy	1201	AACATCATATGACGACATCAGATATTAACAAGAAACAACCCCAAAAATTTCTTCGGGAT	1260
Db	7140	GGCCCCACCAACAAAGGCAACCAATATGACAAATGGGCAATTTAAGCAGTCCCTCCCCCACC	7199
Qy	1261	TCTAGTCCGACAAACCGCCCTCTATATACITTTAGAAAGAAACGATCGATCCCTCTGGAGG	1320
Db	7200	CCCAACCCGACCAACAACATCTTGTATATTTTCAGAAAGAAACGAAGTATCTCTGGAGG	7259
Qy	1321	GAAGCGACATGTTCCCTTTTCTGGATGGGTTAATAAATGCTCCAAATTTGATTTGACCCA	1380
Db	7260	GAAGCGACATGTTTCTTTTCTGACGGGTTAATAATGCTCCAAATTTGATTTGATCCA	7319
Qy	1381	GTTCCAAATAAAAAACAATCTTTGATGAATCTCTTAGTTCTGTGTGCTCGGCTGAGGAA	1440
Db	7320	GTTTCCAAATACAAAGACGATCTTTGATGAATCTTTCTAGTTCTGTGTGCTTCGGCTGAGGAA	7379

Qy	1441	GATCAACATGCCTCCCCCAATATAGTTTAACTTTATCTTATTTCTTCTTAATAATAATGAG	1500
Db	7380	GATCAACATGCCTCCCCCAATATAGTTTAACTTTATCTTATTTCTTCTTAATAATAATGAA	7439
Qy	1501	AACACTGCCTACTCTCGAGAAAATGAGAATGATTGTGATGCAGAGTTTAAGAAATTTGGAGC	1560
Db	7440	AACACTGCCTACTCTCGAGAAAATGAGAATGATTGTGATGCAGAGTTTAAGAAATTTGGAGC	7499
Qy	1561	GTTTCAGGAGGATGACCTGGCCGAGGCGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT	1620
Db	7500	GTTTCAGGAGGATGACCTGGCCGAGGCGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT	7559
Qy	1621	GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAACAAATTTGGTCTCGCAGGTTGAGG	1680
Db	7560	GAAGGACTTTATATCTGCTGGTTTTTAATTAATAAATCAAAACAAATTTGGTCTCGCAGGTTGAGG	7619
Qy	1681	CGTCTAGCCAAATCAAACTGCAAAATCCTTGGAACTCTTATTGTAGAGTCACAACTGTAGGAA	1740
Db	7620	CGTCTAGCCAAATCAAACTGCAAAATCCTTGGAACTCTTATTGTAGAGTCACAACTGTAGGAA	7679
Qy	1741	AGAACATTTCTCTTAATCAATAGACATGCTATTGTGACTTCTTACTCAAGATGGGGAGGA	1800
Db	7680	AGAACATTTCTCTTAATCAATAGACATGCTATTGTGACTTCTTACTCAAGATGGGGAGGA	7739
Qy	1801	ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGGATAGAGACTTGTTCCTCAAAAATATT	1860
Db	7740	ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGGATAGAGACTTGTTCCTCAAAAATATT	7799
Qy	1861	TCAGAGCAAAATGACCAAAATTAAGAAAGGACGACAAAGAGAGGGGACTGGTTGGGGCTCG	1920
Db	7800	TCAGAGCAAAATGACCAAAATTAAGAAAGGACGACAAAGAGAGGGGACTGGTTGGGGCTCG	7859
Qy	1921	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGTCTACTA	1980
Db	7860	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGTCTACTA	7919
Qy	1981	TTATCCATAGCTGCTTTGATTTGCTCTATCCTGTATTTGTTCGTATCTTTAATAATATTC	2040
Db	7920	TTATCCATAGCTGCTTTGATTTGCTCTATCCTGTATTTGTTCGTATCTTTAATAATATATC	7979
Qy	2041	GGATAA 2046	
Db	7980	GGGTAA 7985	
RESULT 11			
ACC71551			
ID	ACC71551 standard; DNA; 7005 BP.		
XX	ACC71551;		
XX	10-JUL-2003 (first entry)		
DT	VRC6702 (pVR1012-x/s Marburg GP (dTM)) plasmid.		
DE	Virucide; vaccine; immune response; Ebola virus; Marburg virus;		
KW	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;		
KW	viral infection; filovirus; circular; cyclic; ds.		
XX	Synthetic.		
OS	WO2003028632-A2.		
PN	10-APR-2003.		
PD	24-SEP-2002; 2002WO-US030251.		
PF	01-OCT-2001; 2001US-0326476P.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PR	Nabel GJ, Yang Z, Sullivan N, Sanchez A;		
PA			
XX			
PI			

DR WPI; 2003-371961/35.
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 188-190; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This vector expresses the Marburg virus glycoprotein
CC without its transmembrane and intracellular domains
XX
SQ Sequence 7005 BP; 1931 A; 1758 C; 1588 G; 1728 T; 0 U; 0 Other;

Query Match 69.3%; Score 1418.8; DB 10; Length 7005;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1618; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
DB 2034 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 2093

QY 61 GTTTTAGAAAATGCTAGTAACAGCCAACTCTCAAGATGTAGATTTCAGTGTGCTCCGGAACT 120
DB 2094 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAACT 2153

QY 121 CTCGAAAGACAGAGATGTTTCATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 180
DB 2154 CTCGAAAGACAGAGATGTTTCATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 2213

QY 181 GATTCCCTTTTGGAGGATCTAAAGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 240
DB 2214 GATTCCCTTTTGGAGGATCTAAAGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 2273

QY 241 GTTGAGTATACGGAAGGAGAGAACCCAAACATGTTTCAATATATAAGTGTAAACAGACCT 300
DB 2274 GTTGAGTATACGGAAGGAGGAGAACCCAAACATGCTTCAATATATAAGTGTAAACAGATCCC 2333

QY 301 TCTGGAATAATCTTCTGCTGATGATCTCCAGTAAATATCCGGATTAACCTTAATATGTA 360
DB 2334 TCTGGAATAATCTTCTGCTGATGATCTCCAGTAAATATCCGGATTAACCTTAATATGTA 2393

QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCCATTTGG 420
DB 2394 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCCATTTATGG 2453

QY 421 GGGGCATTTTCTGATGATGCTGCTCTCAACAATATGTTACCGAGGCAAGGTCTTC 480
DB 2454 GGAGCATTTTCTGATGATGCTGCTCTCAACAATATGTTACCGAGGCAAGGTCTTC 2513

QY 481 ACTGAAGGAATATATAGCAGCTATGATTTGTTAATAGACAGTTTCAAGAATGATTTTTCT 540
DB 2514 ACTGAAGGAATATATAGCAGCTATGATTTGTTAATAGACAGTTTCAAGAATGATTTTTCT 2573

QY 541 AGGCAAGGACAGGTTTATCGTCACATGAATTTGACCTCCACCAATATAATTTGGACAGC 600
DB 2574 CGGCAAGGACAGGTTTATCGTCACATGAATTTGACCTTTCTATATATAATTTGGACAGT 2633

QY 601 AGCAATGAACCGACAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATATCAACTCC 660
DB 601 AGCAATGAACCGACAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATATCAACTCC 660

QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGAGGA 1800
D 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGAGGA 1800
Db 3774 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGAGGA 3833
QY 1801 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGATAGAGACTTGTCCAAAAATATT 1860
D 1801 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGATAGAGACTTGTCCAAAAATATT 1860
Db 3834 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGATAGAGACTTGTCCAAAAATATT 3893
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGTTGGGGTCTG 1920
D 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGTTGGGGTCTG 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGTTGGGGTCTG 3953
QY 1921 GGTGTAATGCTGACATCCGACTGGGGT 1950
D 1921 GGTGTAATGCTGACATCCGACTGGGGT 1950
Db 3954 GGTGTAATGCTGACATCCGACTGGGGT 3983

RESULT 12
ACCT1552
ID ACC71552 standard; DNA; 8256 BP.
AC AC
XX ACC71552;
DT 10-JUL-2003 (first entry)
XX VRC6710 (pAdapt Marburg GP(dTM)) plasmid.
DE Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX Synthetic.
OS
XX WO2003028632-A2.
PN 10-APR-2003.
PD 24-SEP-2002; 2002WO-US030251.
PF 01-OCT-2001; 2001US-0326476P.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
DR New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX Claim 1; Page 190-193; 219pp; English.
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence is an adenovirus shuttle vector,
CC expressing the Marburg virus glycoprotein without its transmembrane and
CC intracellular domains
XX Sequence 8256 BP; 2048 A; 2049 C; 2108 G; 2051 T; 0 U; 0 Other;
SQ

Query Match 69.3%; Score 1418.8; DB 10; Length 8256;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1618; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
D 1431 ATGAAGACCAATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 1431 ATGAAGACCAATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 1490
QY 61 GTTTTAAAGATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAACC 120
D 1491 ATTTTAAAGATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAACC 1550
Db 1491 ATTTTAAAGATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAACC 1550
QY 121 CTCCAAAGACAGAAAGATGTTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
D 1551 CTCCAAAGACAGAAAGATGTTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 1610
Db 1551 CTCCAAAGACAGAAAGATGTTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 1610
QY 181 GATTCCCTTTTGAAGCATCTTAAACGATGGCTTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
D 1611 GATTCCCTTTTGAAGCATCTTAAACGATGGCTTTTTCAGGACAGGTGTTCTCCCAAGAAC 1670
Db 1611 GATTCCCTTTTGAAGCATCTTAAACGATGGCTTTTTCAGGACAGGTGTTCTCCCAAGAAC 1670
QY 241 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 300
D 1671 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 1730
Db 1671 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 1730
QY 301 TCTGGAATAATCTTGTGCTGGATCTCCCGATTAATATCCCGATTAATATTAATGTTAAA 360
D 1731 TCTGGAATAATCTTGTGCTGGATCTCCCGATTAATATCCCGATTAATATTAATGTTAAA 1790
Db 1731 TCTGGAATAATCTTGTGCTGGATCTCCCGATTAATATCCCGATTAATATTAATGTTAAA 1790
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGGGGGATTCCTCCATTTGCG 420
D 1791 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGGGGGATTCCTCCATTTGCG 1850
Db 1791 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGGGGGATTCCTCCATTTGCG 1850
QY 421 GGGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTATCCGAGGCAAGGTCTTC 480
D 1851 GGGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTATCCGAGGCAAGGTCTTC 1910
Db 1851 GGGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTATCCGAGGCAAGGTCTTC 1910
QY 481 ACTGAAGAAATATAGCAGCTATGATTGTTTAAAGACAGTTCACAGAAATGATTTTCT 540
D 1911 ACTGAAGAAATATAGCAGCTATGATTGTTTAAAGACAGTTCACAGAAATGATTTTCT 1970
Db 1911 ACTGAAGAAATATAGCAGCTATGATTGTTTAAAGACAGTTCACAGAAATGATTTTCT 1970
QY 541 AGCAAGGACAGAGGTATTCGTCATGACATGACCTCCACCAATAAATATTTGGAAGAC 600
D 1971 AGCAAGGACAGAGGTATTCGTCATGACATGACCTCCACCAATAAATATTTGGAAGAC 2030
Db 1971 AGCAAGGACAGAGGTATTCGTCATGACATGACCTCCACCAATAAATATTTGGAAGAC 2030
QY 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGATATCAACTCC 660
D 2031 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGATATCAACTCC 2090
Db 2031 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGATATCAACTCC 2090
QY 661 ACAAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACTCCG 720
D 2091 ACAAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACTCCG 2150
Db 2091 ACAAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACTCCG 2150
QY 721 AGCAATCTACTACAAATCTCAAAATTAATCTGTAATCTGGAATCTGGAATCTGGAAT 780
D 2151 AGCAATCTACTACAAATCTCAAAATTAATCTGTAATCTGGAATCTGGAATCTGGAAT 2210
Db 2151 AGCAATCTACTACAAATCTCAAAATTAATCTGTAATCTGGAATCTGGAATCTGGAAT 2210
QY 781 AGCGACATGAGGACCTTATGATTTCGGGCTCAGGATCTGAGAACAGGGGCCCCACACA 840
D 2211 AGCGACATGAGGACCTTATGATTTCGGGCTCAGGATCTGAGAACAGGGGCCCCACACA 2270
Db 2211 AGCGACATGAGGACCTTATGATTTCGGGCTCAGGATCTGAGAACAGGGGCCCCACACA 2270
QY 841 ACTCTTAATGATGATCTGAAACAGAAACAAATGCTGCAACAATATTGTCACCTCTTCACTA 900
D 2271 ACTCTTAATGATGATCTGAAACAGAAACAAATGCTGCAACAATATTGTCACCTCTTCACTA 2330
Db 2271 ACTCTTAATGATGATCTGAAACAGAAACAAATGCTGCAACAATATTGTCACCTCTTCACTA 2330
QY 901 CATCCAGCACCTCACAACATGAGCAAAACAGTACGATCTCCCGACATGCTGTAAT 960
D 2331 CATCCAGCACCTCACAACATGAGCAAAACAGTACGATCTCCCGACATGCTGTAAT 2390
Db 2331 CATCCAGCACCTCACAACATGAGCAAAACAGTACGATCTCCCGACATGCTGTAAT 2390
QY 961 GAGCACATGAAACCGACCCCAACCAACAGCAACAGCTCTCCTCAACAATACTAATACA 1020
D 2391 GAGCACATGAAACCGACCCCAACCAACAGCAACAGCTCTCCTCAACAATACTAATACA 2450
Db 2391 GAGCACATGAAACCGACCCCAACCAACAGCAACAGCTCTCCTCAACAATACTAATACA 2450

Db 421 GGAGCATTTTTTCTGTATGTCGATTCGCTCCACAAATGTATCCGAGGCARAGTCTTC 480
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATGAAGACAGTTTCAAGAAATGATTTTTTCT 540
Db 481 ACTGAAGGAAACATAGCAGCTATGATTTGTTCAATGAAGACAGTGCACAAATGATTTTTCTCG 540
Qy 541 AGGCAAGCAAGGTTATCGTCACATGAATGACCTCCACCAATATATATGACAGC 600
Db 541 CGGCAGAGCAAGGTTACCGTCAATGAACTGACTTCTACTAATAATATTTGACAAAT 600
Qy 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTTGGCATCTCTCCAGAAATACAACTCC 660
Db 601 AGTAACGAGACGCAAAACGAATGACACTGGATGTTTTCGSGCTCTTCAAGAAATACAACTCT 660
Qy 661 ACAAAATCAAAATGATCCCTCATCTCTTTAACTCCATCCCTGCGCCACAGTAATCTCG 720
Db 661 ACAAGAACCAAAATGCTGCTCCGTCACAAATATCTCCACCACTGCGCCACAGCCCGTCCG 720
Qy 721 AGCATTCACTCTACAAATCTCAAAATTAATTAATCTGAACTGAACTGAACTGAACTGAACT 780
Db 721 GAGATCAAACTCAACGACCCCAACTGATGTCACCAAACTCAATATCACGGAACCCCAAGC 780
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAAACAGGGGCCCCACACA 840
Db 781 AGTGATGATGAGGACCTCGCAATCCGGCTCAGGGTCCGGAGAAACGAGAACCCACACA 840
Qy 841 ACTCTTAATGATGATGACAGAAACAAATGCTCAACAAATATTTGTCACCTCTTCACTA 900
Db 841 ACTTCTGATGGGTCAACCAAGCAGGGCTTTTCATCAACAAATGCTCCACCACTCCCTCA 900
Qy 901 CATCAAGCCTCACAACATGACGAAACAGTAGGATCTCTCCGACATGCTGTAATCT 960
Db 901 CAACCAAGCAGCCACAGCAAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT 960
Qy 961 GAGCACAATGGAACCGACCCCAACCAACCAACGACGCTCTCTCAACAAATTAATACA 1020
Db 961 GAATGAGCAAAATTAACAACTGACACACGCTGACACGCTGCTGACCTCATACACTACCA 1020
Qy 1021 ACTCCACCTTAATACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCGCAAC 1080
Db 1021 ATCTCTACTAAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTAACAAAC 1080
Qy 1081 ATCAACCAATGATACACACGCTGTAATGACGAAAGCGAAGCAACCAATGCTCAGTTG 1140
Db 1081 ACCCAATGACACACAGACGACAACTCACTGAAATGAGCAACCAAGTGCCTCCCTCG 1140
Qy 1141 AACCAACTCTAGATCCAAAGAGAAATCCACACAGCAGACAGCAACCAACAGCACAAC 1200
Db 1141 ATAACAACCTGCTCCAAACGGGAATCCCAACAGCAAGCAAGCAACAGCAAGCAAAA 1200
Qy 1201 AACATCATCATGAGGACATCAGATATATACAGCAAAACACCCCAAAATTTCTTCTCCGGAT 1260
Db 1201 GGCCCCGCCAACACGGCACCAAAACAGCAAAATGAGCAATTCACCACTCTCCCTCCACC 1260
Qy 1261 TCTAGTCGCAACACCGCCCTCTATATATATTTTGAAGAAACGATGATGATCTCTGAGG 1320
Db 1261 CCCAGCTCGACTGCAACAACTCTGTATATTTTCAAGAAAGAGCGATGATTTCTTGGAAA 1320
Qy 1321 GAAGGACATGTTCCCTTTCTCGATGGGTTAATGATGCTCCAAATGATTTTGAACCA 1380
Db 1321 GAAGGTGATATTTCCCGGTTTTAGATGGGTTAATAATATCTGAAATTTGATTTGATCCA 1380
Qy 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCTCTAGTTCTGGTGCCTCCGGCTGAGAA 1440
Db 1381 ATCCCAACACAGAAACAACTTTTGAATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Qy 1441 GATCAACATGCTCCCAATATATGATTTAATCTTTATCTTTTCTTAATATATAATGAG 1500
Db 1441 GAACAAACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAAAATGGA 1500
Qy 1501 AACACTGCTACTCTGGAGAAATCAGATATGATGATGATGATGATGATGATGATGATGATG 1560
Db 1501 GATACTGCTACTCTGGGGAACAGAAATGATTTGTGATGATGATGATGATGATGATGATGAT 1560

Qy 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACGTTTTTTGGCCCTGGAAT 1620
Db 1561 GTGCAGAGGACGATTTTGGCGGCGAGGCTTAGCTGGATACCATTTTTTTGGCCCTGGAATC 1620
Qy 1621 GAAGGACTTTTACACTGCTGTTTTTAAATTAATAAATCAAAACAAATTTGGTCTGCAAGGTTGAGG 1680
Db 1621 GAAGGACTTTTACTGCTGCGGTTTAAATCAAAATCAAGAAATTTAGTTTGTAGTTGAGG 1680
Qy 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTTTATTTGAGAGTCACAACTGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTTTGTTAAGGCTCACAACCGAGGAA 1740
Qy 1741 AGACATTTCTTAAATCAATAGACATGCTATTGAACTTTCTACTCACAAGATGGGAGGA 1800
Db 1741 AGGACATTTTCTTAAATCAATAGACATGCAATTTGCTTACGAGGTGGGCGGA 1800
Qy 1801 ACATGCAAGTGTGTGACCTGATTGTTGTCATCGGATAGAAGACTTGTCCAAAAATAT 1860
Db 1801 ACATGCAAGTGTGTGACCTGATTGTTGTCATAGTAAGATAGAGATCTATCTAAAAATATC 1860
Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGCTCTA 1920
Qy 1921 GGTGTTAAATGTTGGACATCGGACTGGGTGTTCTTACTACTTGGGCAATTTTGGCTACTA 1980
Db 1921 GGTGCAAAATGTTGGACATCGGACTGGGTGTTCTTACTACTTGGGCAATTTTGGGCAAT 1980
Qy 1981 TTATCCATAGCTGTTGATTTGCTCTATCTGTTATTTGCTATCTTTTACTAAATATATC 2040
Db 1981 TTATCTATAGCTGTTGATTTGCTCTGTTCTGTTATCTGCTATCTTCTTACTAAATATATC 2040
Qy 2041 GGATAA 2046
Db 2041 GGATGA 2046

RESULT 14

ACCT1562
ID ACC71562 standard; DNA; 6902 BP.

XX ACC71562;

XX 10-JUL-2003 (first entry)

XX VRC6703 (pVR1012x/s Marburgdelta TM/h (codon optimised)) plasmid.

XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.

XX Synthetic.

XX WO2003028632-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030251.

XX 01-OCT-2001; 2001US-0326476P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;

XX WPI; 2003-371961/35.

XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.

XX Claim 1; Page 213-215; 219pp; English.

XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX

Query Match 40.8%; Score 834; DB 10; Length 6902;
Best Local Similarity 64.3%; Pred. No. 1.3e-231;
Matches 1251; Conservative 0; Mismatches 695; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
DB |||||
QY 1923 ATGAAGACCATCTGCTTTCATCAGCCTGATCTCTGATCCAGGGCATCAAGACCTTGCCC 1982
DB |||||
QY 61 GTTTTGAATAATTGCTAGTAACAGCCCACTCAAGATGTAGATTTCAGTGTCTCGGAACC 120
DB |||||
QY 1983 ATCTGGAGATGCCAGCAACACACGCCCCCAAGACGTGGACGGTGTGACGGGCACC 2042
DB |||||
QY 121 CTCAAAAGACAGAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
DB |||||
QY 2043 CTGCAGAGACCGAGAGCTGCACCTGATGGCTTCACTCCCTGAGCGGCAGAGTGGCC 2102
DB |||||
QY 191 GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTCTCCCAAGAAC 240
DB |||||
QY 2103 GACAGCCCTCTGGAGGCCAGCAAGAGTGGGCTTTTCAGGACCGGCGTGGCCCTCCCAAGAAC 2162
DB |||||
QY 241 GTTGAGTATACGAGGAGAGAGAGCCAAACATGTTACATATATAGTGTACAGACCCCT 300
DB |||||
QY 2163 GTGGAGTACACGAGGGCGAGGAGGCCAAGACCTGCTACATCATCAGCGTGACCGACCCC 2222
DB |||||
QY 301 TCTGAAAATCTTGCTGCTGATCCTCCAGTAAATATATCCGGATTTACCTAAATGTAAA 360
DB |||||
QY 2223 AGCGCAAGAGCGCTGCTGTGGACCTCCCAACATCATCAGGACTACCTTAAGTGCAAG 2282
DB |||||
QY 361 ACTGTTTCATCATATTCAAGGTCAAAACCTCATGACAGGGGATGGCCCTCATTTGTGG 420
DB |||||
QY 2283 ACCATCCACCATCCAGGGCCAGAACCTTCACGCCACAGGGCATCGCCCTGCACCTGTGG 2342
DB |||||
QY 421 GGGCATTTTCTGTATGATCGCGTGTGCTTACAAATGTATCCGAGGCAAGGTCTTC 480
DB |||||
QY 2343 GCGCCTTCTCTGTACAGAGGATCGCAGGACCAACATATGACAGGGCAGGGTGTTC 2402
DB |||||
QY 481 ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTTCACAGAAATGATTTTTCT 540
DB |||||
QY 2403 ACCGAGGCAACATCGCGCCATGATCGTTAACAGACCGTGCAAGATGATCTTCAGC 2462
DB |||||
QY 541 AGGCAAGGACAGGTTATGTCATATGAACTTTGACCTCCACCAATATAATTTGGACAGC 600
DB |||||
QY 2463 AGGCAGGGCCAGGGCTACAGGCACATGAACCTTGACCAAGCAACCAAGTACTTGGACCAAGC 2522
DB |||||
QY 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
DB |||||
QY 2523 AACACCGGACCCAGCAACCAACACCGGCTGCTTCGGCGCCCTCAGAGGTATCAACAGC 2582
DB |||||
QY 661 ACAAACAATCAAAATGCGCTCCATCTCTTAAACCTCCATCTCCCTGCCACACAGTAACTCCG 720
DB |||||
QY 2583 ACCAAGACCGACCTGGCGCCCGCAGCAAGATCCCGACCGCCCTGCCCCACCGCAGGCC 2642
DB |||||
QY 721 AGCATTTCACTTCAAAATCTCAAAATTAATATCTGCTAAATCTGGAACCTATGAAACCAAGT 780
DB |||||

Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;

2643 GAGATCAAGCCACCAGACACCCCAACGACCGCACACCTCTGAACACCAACGACCCCAAC 2702

781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840

2703 AACGACGAGGAGACCTGATCACCAGCGGACGCGGCGGAGGAGGAGGAGGAGGAGGAGG 2762

841 ACTCTTAATGTAGTACTGTAAACAGAAACAATGTCACAAATATTGTTCACCTCTTCACTA 900

2763 ACCAGCAGCGCGTGACCAAGCGGCTGAGCAGCACCATGCGCTCTCTACCCCTAGCCCT 2822

901 CATCAAGCACTCACAACATGACCAAAACAGTACGAATCTTCCGACATGCTGTAAT 960

2823 CAGCCCAAGCCTCTAGCAGGAGGGCAACAAACACCAACGACGAGGACCGCTGAC 2882

961 GAGCACAAATGGAAACCGACCCCAACACAAACAGCAAGCTCTCTCAACAATACTAATA 1020

2883 GAGCCCAAGACCAACACCCAGCCAGCCAGCATGCTCTCTCAACACACACCCGCC 2942

1021 ACTCCCACTTAAACACTCTCAAGTACAACTCTAGTACTCTTCCCTCTCAACCCGCAAC 1080

2943 ATCAGCACCAACACACCCAGCAAGCAACATTTTACGACCTCTGAGCGTGAGCGCTGAGA 3002

1081 ATCACCATAATGATACACACGTGAATAGCAGAAAGCGCAACAAACCAATGCTCAGTTG 1140

3003 ACCACCACTACGACACCCAGAGCACCGCCACCGAGAACGAGCAGACGAGCGCCCTAGC 3062

1141 AACACAACTCTAGATCCAAACAGAAAATCCCAACAGACAGCAAGACACCAACAGCAAC 1200

3063 AAGACCACTGCTCTCCACCGGCAACCTGACCAACCGCCAGAGACCAACCAACACCAAG 3122

1201 AACATCATATGACGACATCATGATATTAACAGCAAAACACCCCAACAAATCTTCTCCGAT 1260

3123 GGGCCCAACCAACCGCCCTTAAATGATCAACACCGCCACCTGACACGCGCCACCCACCC 3182

1261 TCTAGTCGACCAACCGCCCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320

3183 CCCAACCCCAACCCAGCACCTGGTGTACTTCAGGAAGAAAGAGAGGAGCATCTCTGGAGG 3242

1321 GAAGCGACATGTTTCCCTTTTCTGGATGGGTAAATTAATGTCTCAATTTGATTTTGACCCA 1380

3243 GAGGGCGATGTTTCCCTTCTCGACCGCTGATCAACGCCCTATCGACTTCGACCCC 3302

1381 GTTCCAAATACAAAACAACTTTTGTATGAAATCTCTCTAGTCTGTGTCCTCGCGTGAGAA 1440

3303 GTGCCCAACACCAAGACCATCTTCGACGAGACGACGACGAGCGCGCGCCGCGAGGAG 3362

1441 GATCAACATGCTCTCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATAATAATAG 1500

3363 GACGACGACGCGCCCAACATCAGCCTGACCTGACCTGAGCTACTTCCCAACATCAACGAG 3422

1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560

3423 AACACCGCTTACAGCGCGGAGAACGAGAACGACTCGACGCGGAGCTGAGGATCTGGAGC 3482

1561 GTTCAGGAGGATGACTGTGGCGCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620

3483 GTGCAGGAGGACGACTGTGGCGCGCGCTGAGCTGGATTCTCTTCTCGCGCCCGGCTC 3542

1621 GAAAGCTTTACACTGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1680

3543 GAGGGCTGTACACCGCGCGCTGATCAAGAAACAGAAACACCTGGTGTGAGGCTGAGG 3602

1681 CGTCTAGCAATCAAACTGCGCAAACTCTTTTGGAACTCTTTTGGAGAGTCAACACTGAGGAA 1740

3603 AGGCTGGCAACACGACCGCCCAAGCTGAGCTGCTGCTGAGGGTGACCCACCGAGGAG 3662

1741 AGAACAATCTCTTAATCAATAGACATGCTATTGATTTCTTACTCAAGATGGGAGGA 1800

3663 AGGACCTTCAGCCTGATCAACAGGACCGCATTCGACTTCTGCTGACAGGTTGGGCGCGC 3722

1801 ACATGCAAGTCTTGGACCTGATTTGTGATCGGATAGAGACTTGTCCAAAAATATT 1860

3723 ACCTGCAAGGTGCTGGGGCCCGGACTGCTGCTCGGATCGAGGACCTGAGCAGGAACATC 3782

Db 1261 CCCAGCTCGACTGCGACAAACATCTTGTATATTTCAGAGAAAGCGATCG 1308

Search completed: July 31, 2005, 16:57:29
Job time : 1077.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6679.18 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatattttct.....ttactaaatatcggataa 2046

Scoring table: IDENTITY_NUC
Gapex 10.0, Gapex 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	85.6	4.2	895	9	CNS0071A
C 2	80.8	3.9	1101	9	CNS00FXE
C 3	80.4	3.9	922	9	CNS0073W
C 4	79.8	3.9	909	9	CNS00JTL
C 5	74.8	3.7	884	9	CNS006UO
C 6	72	3.5	1101	9	CNS00LOO
C 7	69	3.4	1101	9	CNS0100X
C 8	67.4	3.3	1225	9	CNS0166K
C 9	66.2	3.2	1101	9	CNS0172T
C 10	66	3.2	1101	9	CNS00LT2
C 11	65.8	3.2	1101	9	CNS016K3
C 12	65.4	3.2	769	8	B2579780
C 13	64.4	3.1	1797	9	AG390561
C 14	64.2	3.1	1179	9	AG117870
C 15	64	3.1	1052	6	CD389522
C 16	63.8	3.1	939	9	CNS00CNG
C 17	63.6	3.1	860	9	CNS018FL
C 18	63	3.0	937	9	CNS006ST
C 19	62.4	3.0	861	9	CNS0075A
C 20	62.2	3.0	935	9	AG127540
C 21	61.6	3.0	1037	9	AG379951
C 22	61	3.0	1325	9	CG754898
C 23	60.8	3.0	905	9	AG129358
C 24	60.6	3.0	926	8	AQ782232

25	60.4	3.0	1095	9	CNS04428	AL274733 Tetraodon
C 26	59.6	2.9	902	9	CNS006QP	AL065804 Drosophil
C 27	59.6	2.9	1052	9	CNS00DIP	AL060519 Drosophil
C 28	59.4	2.9	1101	9	CNS017WI	AL108588 Drosophil
C 29	59.4	2.9	1188	9	CG747780	CG747780 P041-2-HO
C 30	59.2	2.9	1974	9	CL467385	CL467385 SAIL 1269
C 31	58.6	2.9	1257	9	AG289995	AG289995 Mus muscu
C 32	58.6	2.9	1499	9	AG371052	AG371052 Mus muscu
C 33	58.4	2.9	1233	8	BZ696963	BZ696963 SP_Ba009
C 34	58.4	2.9	1779	9	AG390408	AG390408 Mus muscu
C 35	58	2.8	891	5	BU530351	BU530351 AGENCOURT
C 36	57.8	2.8	1713	9	CG754301	CG754301 P049-3-GO
C 37	57.6	2.8	779	9	AG519641	AG519641 Mus muscu
C 38	57.4	2.8	768	9	AG599644	AG599644 Mus muscu
C 39	57.4	2.8	1461	9	CL647640	CL647640 CH213-143
C 40	57.4	2.8	2050	9	AG365337	AG365337 Mus muscu
C 41	57.2	2.8	1226	9	CG757449	CG757449 P052-4-AO
C 42	57	2.8	1144	9	AG126238	AG126238 Pan trogl
C 43	57	2.8	1185	5	BQ650581	BQ650581 AGENCOURT
C 44	57	2.8	1759	9	CG754278	CG754278 P049-3-F0
C 45	56.8	2.8	513	9	CNS02CMP	AL191328 Tetraodon

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL066286.1 GI:4945153
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 895)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..895
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/clone_lib="RPI-98"
/note="end : TET3"

ORIGIN

Query Match 4.2%; Score 85.6; DB 9; Length 895;
Best Local Similarity 22.5%; Pred. No. 1.9e-12;

```
Matches 96; Conservative 165; Mismatches 164; Indels 1; Gaps 1;
QY 855 CACTGAACAGAAACAATCGTCAACAATATTGTCCACTCTTCACTATCAATCAAGCACCTC 914
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
889 CCCMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 830
QY 915 ACAACATGAGCAAAACAGTGAATTCCTTCCGACATCTGTAACTGAGCACAATGGAA 974
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
829 MMAMMCCCMCMCMCAAAAMCMAMMMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 770
QY 975 CGACCAACACACACAGCAACGCTCTCTCAACATTAATACTAATACTCCACCTATA 1034
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
769 MMAMCMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 710
QY 1035 CACTCTCAAGTACAACCTCAGTACTCTCTCCCTCCAAACCCGACACATCAACAATGA 1094
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
709 MCAAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 650
QY 1095 TACAACAGTGAATGAGCAAGAAAGCAACAAACCAATCTCAGTTGAAACAACCTCTAGA 1154
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
649 MAMCCAMMCAAMM-AMCMCMCMCCCAAMCMCAACCCMMCMCMAMAMAMAMAMAMAM 591
QY 1155 TCCACAGAAATCCCAACAGGACAGACAGCAACCAACCAACATCATCATGAC 1214
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
590 MMAMCMCMAMMAMCMAMMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 531
QY 1215 GACATCAGATATAACAAGCAACACCCACCAAAATTTCTTCGGATTCTAGTCCGACAAC 1274
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
530 MMAMAMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 471
QY 1275 CGCCC 1280
Db |:::|
470 CACCM 465

RESULT 2
CNS00FXE/c 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL071370.1 GI:4951210
VERSION AL071370
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
DIRECT SUBMISSION Drosophila melanogaster (fruit fly)
SUBMITTED (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR32C19"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 3.9%; Score 80.8; DB 9; Length 1101;
Best Local Similarity 25.7%; Pred. No. 4.5e-11;
Matches 105; Conservative 130; Mismatches 174; Indels 0; Gaps 0;

QY 906 AAGCACCCTCAACATGAGCAAAACAGTACGAATCTTCCGACATGCTGTAACTGACGA 965
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1101 AAAMAMATAAMMAHAYACAAAAMAMAMMMMAACAAMMAMAMAMAMAMAMMCMCMMA 1042
QY 966 CAATGGAAACGCGACCAACACACACAGCAACGCTCTCAACAATACTAATAACAATCC 1025
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1041 MAMMAMACAGMHCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMMA 982
QY 1026 CACCTATAACACTCTCAAGTACAAACCTCAGTACTCTTCCCTCCAAACCCGACACATC 1085
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
981 MAMCMAMMMMNCAAMCMAMAMAMAMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 922
QY 1086 CAATAATGATACACAACGTAAGTACAGAGAAAGGACAAACCAATGCTCAGTTGAAAC 1145
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
921 MACCCATACATMMTMMCMMAAAMAMAMMAAAMAMAMAMAMAMAMAMAMAMAMAMAAA 862
QY 1146 AACTCTAGATCCCAAGAAATCCCAACAGGACAGACACCAACAGACACCAACCAAT 1205
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
861 AAMAMAMAMAMAMAMAMAMAMAMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 802
QY 1206 CATCATGACGACATCAGATATAACAAGCAACACCCACCAAAATTTCTTCGGATTCTAG 1265
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
801 MCCAMMAMCMCMCSAGMAMCMCMMAAMAMAMAMAMAMAMCMCMCMCMCMCMCMCMCM 742
QY 1266 TCCGACAAACCGCCCTCTCTATATATCTTTAGAAAGAAAGATCGATCTCTC 1314
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
741 CMCMAMAMCCCMCMACCCCCCTCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 693

RESULT 3
CNS0073W/c 922 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066784
VERSION AL066784.1 GI:4945247
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 922)
AUTHORS Genoscope.
DIRECT SUBMISSION Drosophila melanogaster (fruit fly)
SUBMITTED (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
```



```

QY 1181 AAGACACCAAGCAGCAACCAACATCATCATGACGACATGATATATAACAGCAACACC 1240
Db 767 ACCCCCCHCCCCCMAAAMAAAATAAAKTYMAAACAACAAMAAAYCMCCGRRARVMSSS 708
QY 1241 CCAGAAA 1247
Db 707 GVSVVV 701

RESULT 9
CNS017ZT/c 1101 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION BACN37N10 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108707
VERSION AL108707.1 GI:5629011
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CDPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37N10"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 3.2%; Score 66.2; DB 9; Length 1101;
Best Local Similarity 31.4%; Pred. No. 6.3e-07;
Matches 143; Conservative 113; Mismatches 194; Indels 6; Gaps 2;

QY 887 CCATCTCTTCACTACATCAAGCACCTCACACATGACGACAAACAGTAGCATCTTCCC 946
Db 1095 CCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCC 1036

QY 947 GACATGCTGTAACTGAGCAGCAATGGAACCGACCAACCAACCAACGAGCTCTCTCA 1006
Db 1035 CAAWACACWACWACWACWACWACWACWACWACWACWACWACWACWACWACWACWACW 976

QY 1007 ACAATACTAATACTACCTCCACCTTAAACACTCTCAAGTACAACCTCAGTACTCTTCCC 1066
Db 975 MWACAMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCC 916

QY 1067 CTCGACCCGACATGATGACCAATATGATACACACGTGAGTACGAGAAAGCGACAAA 1126
Db 915 AMCCAAMAWACACAWACACACACCCCC---CNCMAWAAAAAAMMAWAAAAAACHCC 860

QY 1127 CCAATGCTCAGTTGAACAACAACCTCTAGATCCAACAGAAAAATCCCAACAGACAGACA 1186
Db 859 CCACWACWACWACWACWACWACWACWACWACWACWACWACWACWACWACWACWACW 800

QY 1187 CCAACAGCACACCAACATCATCATGACGACATCATGATATATAACAGCA--AACACCCAC 1244

```

```

Db 799 CCMWCAACMCMYAMCMCAWCHYCTTTTTTTTTTACWYCMAAAHMAYAYATTACTYIC 740
QY 1245 AAAATTCTTCCGATTCTAGTCGCAACACCGCCCTCTATATACTTTTAGAAGAAACG 1304
Db 739 AAAMWYAYYYMYMYCTCTHMYATACMYCCCYAYHYWYTWYTWCMCTWTHAMWMTWY 680
QY 1305 ATCATCTCTGAGGAGGAGGACATGTTCCCTTT 1340
Db 679 TYATCMCAAMCTMMCAAWTCATCMCCCTCCCTTY 644

RESULT 10
CNS00LT2/c 1101 bp DNA linear GSS 14-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC
DEFINITION BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL078714
VERSION AL078714.1 GI:5102004
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 3.2%; Score 66; DB 9; Length 1101;
Best Local Similarity 13.1%; Pred. No. 7.1e-07;
Matches 53; Conservative 194; Mismatches 157; Indels 0; Gaps 0;

QY 835 CACAACTCTTAATGTAGTCATCGAAGACAGAAACAATCGTCAACAATATTGTCCACTCCT 894
Db 1092 MMWMMCMWHTAAMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 1033

QY 895 TCATCATCTCAACACCTCACACATGACGACAAACAGTAGCATCTCTCCGACATCCT 954
Db 1032 MWMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 973

QY 955 GTAATGAGCAATGGAACCGACCGACCAACACACACAGCAACGCTCCTCAACAATACT 1014
Db 972 TTYMMCMWMTHTTMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMMWMM 913

```

[illegible]

```
QY 1035 CACTTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGCAACATCACCAATAATGA 1094
|||
D 351 CACAACAAACACACACAAACACACCTCAACACCAACAAACACACCAACCCACCC 410
|||
QY 1095 TACACACGTGAATAGCAGAGAAAGGAACAAACCAATCTCTAGTTGAACACAACTCTAGA 1154
|||
D 411 AACACACACACACGAGAAACAAAAAACACCAACACAAACCAACCAACCCCA 470
|||
QY 1155 TCCACAGAAATCCACACACAGACAGACACCAACAGCAGCAACCAACATCATCATGAC 1214
|||
D 471 CACACAAACCAACACACACCCACACACACCAACCAACCAACCAACCAACCA 530
|||
QY 1215 GACATCAGATATAACAGCAACACCCCAACCA 1247
|||
D 531 AAAACATAACACACACACACGACGACCA 563
|||

RESULT 13
AG390561/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-207E07.TJ, genomic survey
sequence.
ACCESSION AG390561
VERSION AG390561.1 GI:48007799
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS BAC end Sequences of Library MSMg01
TITLE Unpublished
JOURNAL 2 (bases 1 to 1797)
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
AUTHORS Direct Submission
TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukuba, Ibaraki, Japan 305-0856, Japan
(E-mail:hattori@riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI.
R.Site 2 : EcoRI.
FEATURES
source
1. 1797
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-207E07.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 3.1%; Score 64.4; DB 9; Length 1797;
Best Local Similarity 47.8%; Pred. No. 2,3e-06;
Matches 196; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

QY 822 AGAACAGGGCCCAACCACTCTTAATGTAGTCTGACGAGAAACAACTCGTCAACAT 881
|||
```

```
D 1531 AANCAACAACAACAACAACAATCAAAACAACAACAACAACAACAACAACAACAACA 1472
|||
QY 882 ATTGTCACCTCTTCACTACATCCAGCAGCCTCTCAACAACATGAGCAAAACAGTAGCAATCC 941
|||
D 1471 ACNACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1412
|||
QY 942 TTCCCGACATGCTGTAATGAGCAATGGAACCGACCCCAACAACAACAACAACAACAACGCT 1001
|||
D 1411 CACNACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1352
|||
QY 1002 CCTCAACAATTAATCAACTCCCACTTAACTCACTCTCAAGTACACCTCTCAGTACTCTCC 1061
|||
D 1351 CCACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1292
|||
QY 1062 TTCCCTCCCAACCGCAACATCACTAATATGATACACAACCAACCAACAACAACAACAACA 1121
|||
D 1291 AACACACACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1233
|||
QY 1122 ACAACAATGCTGATGGAACAACAATCTAGATCCACAGAAATCCCAACCAACGAGACA 1181
|||
D 1232 AAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1173
|||
QY 1182 AGACACCAACAGCAGCAACCAACATCATCATGACGACATCAGATATAACAA 1231
|||
D 1172 CAACACACCAAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1123
|||

RESULT 14
AG117870/c
LOCUS Pan troglodytes DNA, clone: PTB-125N11.R, genomic survey sequence.
ACCESSION AG117870
VERSION AG117870.1 GI:16738389
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1179)
REFERENCE Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukuba, Ibaraki, Japan 305-0856, Japan
(E-mail:chimbeg@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI.
R.Site 2 : SacI.
FEATURES
source
1. 1179
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-125N11.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 3.1%; Score 64.2; DB 9; Length 1179;
```

Best Local Similarity	47.08;	Pred. No.	2.4e-06;
Matches	214;	Conservative	0; Mismatches 239; Indels 2; Gaps 2;
Qy	822	AGACAGGGGCCCCACACAACCTCTTAATGTAGTCAGTGAACAGAAAACAATCGTCAAACAAT	881
Db	564	AAACCANACACCTTAACACTCAGCTCAACAACACACACACACCCCAAAACACTCATAT	505
Qy	882	ATTGTCCAGCTCTTCATACATCCAAGACCTCAACAATGAGCAAAACAGTAGCAATCC	941
Db	504	ACCACACATACAAACCGCGCACCCAGCCACTTACCACGCCAACCAANTACCACACAAAACA	445
Qy	942	TTCOCGACATGCTGTAATCTGAGACAAATGGAACGCACCAACACACAAACCCAGAACGCT	1001
Db	444	ACCAANNNTCTTNACNCMAACAATAACCAACACACACACACNNNAATTAACNCCC	385
Qy	1002	CCTCAACAANTAATAACAATCCCACTCTATPAACTCTCAAGTACAACCTCAGTACTCC	1061
Db	384	ACACACACCCACCCCCGCCCCCNCCCAANAACAAAAACAAAAACACACACANC	325
Qy	1062	TTCCTCCTCCA-ACCGGCAACATACCAATATATGATACAAACGTGTGAATAGCAGAAGCG	1120
Db	324	CACNCNNANACAAAAAANCANRANNAACAAAAACAAAACCCACACCCACAAAAAACA	265
Qy	1121	AACAAACCAATGCTCAGTTGAAACACAACCTCTAGATCCAAACAGRAAATCCCAACACAGAC	1180
Db	264	AACACACCAACACACA-ACACACACAAACACCAACCCCAACAAACACCAACCCACAA	206
Qy	1181	AAGACACCAACAGACCAACCAATCATCTATGAGCATCAGATATATAACAGCAAAACACC	1240
Db	205	ACAANCCCAACAANACACAAACCCCAACAACAACAAAAAANAACANBACAC	146
Qy	1241	CCACAANTTCTTCOGGATTTCTAGTCGCAACCC	1275
Db	145	AAACACACACACANACCAANCAAAAGACCCACC	111

RESULT 15	CD389522	1052 bp	mRNA	linear	EST 30-MAY-2003
LOCUS	CD389522				
DEFINITION	ACENOCOURT_14345368 NIH_MGC_173				Homo sapiens cDNA 5', mRNA sequence.
ACCESSION	CD389522				
VERSION	CD389522.1				GI:31228314
KEYWORDS	EST.				
SOURCE					Homo sapiens (human)

```

FEATURES
source
1. .1052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem
cells"
nrgn quality sequence stop: 204.
location/Qualifiers

```

```

/lab_host="DH10B_TonaA"
/clone_lib="NIH_MGC_173"
/note="Vector: pDONR201; Site 1: attP1;
L1:R PRIME - oligo df; METHOD - full-length enriched;
L1:R PROVIDER - Bradfield"

```

ORIGIN	Query Match	3.1%;	Score 64;	DB 6;	Length 1052;
	Best Local Similarity	42.7%;	Pred. No. 2.6e-06;		
	Matches 193;	Conservative 0;	Mismatches 259;	Indels 0;	Gaps 0;
Qy	863	AGAAACAATGGTCAACAATATTGTGTCACTCTCTTCACTATATCCAGCACCTCAACATG	922		
Db	282	AGNAAAAACCTGGAAAGGCCAGGTCCCTGCGCCCCCNNNNNNNNNNNNNNN	341		
Qy	923	AGCAAAACAGTAGCGAATCTCTCCGACATGCTGTAACCTGAGCACAAATGGAACCGACCCAA	982		
Db	342	NN	401		
Qy	983	CAACACAACGACGACGCTCTCAACAATACTAATAACAACCTCCACGCTATAACACTCTCA	1042		
Db	402	CCAAACCCCAACAACCCCAACAACAACCCCCCCCCCCCCCCCCCCCAACCAACCAC	461		
Qy	1043	AGTACAACCTCGATGACTCTCTTCCCTTCCAAACCGCAACATCAACCAATATGATACAAC	1102		
Db	462	ACAAACAACCAACCCCAACCAACNACCACACACCANACCCTCCCAACCAACACCCCAACC	521		
Qy	1103	GTGAACCTAGCAGNAGCGAACAAACCAATGCTCAGTTGAAACAACAATCTAGATCCAACAG	1162		
Db	522	CAAAAAAAACACACACCAACAACCAACCCCAACACCCCAACCAACCAACCAAAAAAACAC	581		
Qy	1163	AAATCCCAACGACGACGAAGACACCAACGACGACCAACCAATCATCATGACGACATCAG	1222		
Db	582	ACAAACCAANAAAAAANAANAACCAACACCAACCAANAAAAAANAANAACCAACCAACCA	641		
Qy	1223	ATATAACAAGCAACACCCCAACAAATTTCTTCTCGGATCTTAGTCGCAACACCCGCCCTC	1282		
Db	642	AAACACCAACCCCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACN	701		
Qy	1283	CTATATATCTTTGAAAGAACGATCGATCCTC	1314		
Db	702	CCCCCAACCCCAACCAAAAAACNACACACCCAC	733		

Search completed: August 1, 2005, 10:09:12
Job time : 6682.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 327.71 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatattttct.....ttactaaatatatcgataa 2046

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1822	89.1	2247	3	US-08-760-615-5
2	1514.8	74.0	2164	3	US-08-760-615-3
3	1514.8	74.0	2164	3	US-08-760-615-3
4	134	6.5	2172	3	US-09-336-910A-1
5	134	6.5	2298	4	US-08-760-615-1
6	54	2.6	5163	3	US-09-650-086A-1
7	54	2.6	5163	3	US-08-700-651-1
8	54	2.6	5163	3	US-08-928-361B-4
9	54	2.6	5163	3	US-08-928-361B-4
10	54	2.6	5318	3	US-08-700-651-2
11	54	2.6	5318	4	US-09-588-995A-3
12	48.8	2.4	631	4	US-09-270-767-2848
13	48.8	2.4	631	4	US-09-270-767-18130
14	48	2.3	5511	3	US-08-928-361B-2
15	48	2.3	5511	4	US-09-588-995A-2
16	48	2.3	7334	3	US-08-928-361B-1
17	48	2.3	7334	4	US-09-588-995A-1
18	47	2.3	767677	4	US-09-949-016-12147
19	47	2.3	767677	4	US-09-949-016-17361
20	43.6	2.1	832	4	US-09-621-976-2813
21	43.4	2.1	1141	4	US-09-806-708B-22
22	43.2	2.1	700	3	US-09-236-097-11
23	43	2.1	152132	4	US-09-949-016-13845
24	43	2.1	152145	4	US-09-949-016-12371
25	43	2.1	325034	4	US-09-949-016-14957
26	43	2.1	389504	4	US-09-949-016-11774
27	42.4	2.1	49559	4	US-09-949-016-12892

28	42.4	2.1	49559	4	US-09-949-016-17267
29	41.4	2.0	1968	4	US-09-248-796A-1729
30	40.8	2.0	7218	1	US-08-432-463-14
31	40.8	2.0	96987	4	US-09-949-016-14429
32	40.6	2.0	55195	4	US-09-949-016-15854
33	40.4	2.0	582	4	US-09-787-292-3
34	40.2	2.0	601	4	US-09-949-016-21189
35	40.2	2.0	601	4	US-09-949-016-158630
36	40.2	2.0	1086	1	US-08-415-751-27
37	40.2	2.0	1086	1	US-08-415-751-28
38	40.2	2.0	17723	4	US-09-949-016-15832
39	39.8	1.9	1397	4	US-09-216-393B-343
40	39.8	1.9	1397	4	US-09-216-393B-345
41	39.8	1.9	60990	4	US-09-949-016-14080
42	39.6	1.9	867	4	US-09-216-393B-340
43	39.6	1.9	867	4	US-09-216-393B-342
44	39.6	1.9	938	3	US-09-078-294-16
45	39.4	1.9	601	4	US-09-949-016-144692

ALIGNMENTS

RESULT 1
US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Ravn
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..2142
; US-08-760-615-5

Query Match 89.1%; Score 1822; DB 3; Length 2247; Best Local Similarity 93.2%; Pred. No. 0; Matches 1906; Conservative 0; Mismatches 140; Indels 0; Gaps 0;									
Qy	1	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT	60						
Db	97	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT	156						
Qy	61	GTGTTAGAAATGCTAGTAAACAGCCAACTCAAGATGTAGATTCAGTGTGCTCCGGAACC	120						
Db	157	GTGTTAGAAATGCTAGTAAACAGCCAACTCAAGATGTAGATTCAGTGTGCTCCGGAACC	216						
Qy	121	CTCCAAAGACAGAAAGTGTTCATCTGATGGATTTTACATGAGTGGCCAAAAGTTGCT	180						
Db	217	CTCCAAAGACAGAAAGTGTTCATCTGATGGATTTTACATGAGTGGCCAAAAGTTGCT	276						
Qy	181	GATTTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTCTCCCAAGAAC	240						
Db	277	GATTTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTCTCCCAAGAAC	336						
Qy	241	GTGAGTATACGGAAGGAGAAAGCCAAACATGTTACAATATAAGTGTAAACAGACCT	300						
Db	337	GTGAGTATACGGAAGGAGAAAGCCAAACATGTTACAATATAAGTGTAAACAGACCT	396						
Qy	301	TCTGAAAAATCCTTGCTGCTGGATCCTCCAGTAAATATCCGCAATACCCCTAAATGTAAA	360						
Db	397	TCTGAAAAATCCTTGCTGCTGGATCCTCCAGTAAATATCCGCAATACCCCTAAATGTAAA	456						
Qy	361	ACTGTTTCATATATCAAGGTCAAAACCTCATGCACAGGGATGGCCCTCCATTTGTGG	420						
Db	457	ACTGTTTCATATATCAAGGTCAAAACCTCATGCACAGGGATGGCCCTCCATTTGTGG	516						
Qy	421	GGGGCATTTTCTTGATGATCGGTTCCTCTACAAACATGTACCGAGGCAAGTCTTC	480						
Db	517	GGGGCATTTTCTTGATGATCGGTTCCTCTACAAACATGTACCGAGGCAAGTCTTC	576						
Qy	481	ACTGAAGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTCACAGAAATGATTTTTCT	540						
Db	577	ACTGAAGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTCACAGAAATGATTTTTCT	636						
Qy	541	AGGCNAGCAAGGTATTCGTACATGAATCTTGACCTCCACCAATATAATTTGGACAAGC	600						
Db	637	AGGCNAGCAAGGTATTCGTACATGAATCTTGACCTCCACCAATATAATTTGGACAAGC	696						
Qy	601	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCAAAGAAATACAACTCC	660						
Db	697	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCNAGNATACAACTCC	756						
Qy	661	ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACTCC	720						
Db	757	ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACTCC	816						
Qy	721	AGCATTCACCTACAAATCTCAAAATTAATCTGCTAAATCTGGAATCTGAACCCAAAGT	780						
Db	817	AGCATTCACCTACAAATCTCAAAATTAATCTGCTAAATCTGGAATCTGAACCCAAAGT	876						
Qy	781	AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAAACAGGGGCCACACA	840						
Db	877	AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAAACAGGGGCCACACA	936						
Qy	841	ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTTGCCACTCTTCACTA	900						
Db	937	ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTTGCCACTCTTCACTA	996						
Qy	901	CATCCAAGCACCTCACAACATGAGCAAAAACAGTACGAATCCTTCCCGACATGCTGTAACT	960						
Db	997	CATCCAAGCACCTCACAACATGAGCAAAAACAGTACGAATCCTTCCCGACATGCTGTAACT	1056						
Qy	961	GAGCACAATGGAAACCGACCCCAACACACACAGCAACGCTCTCTCAACAATACTAATACA	1020						
Db	1057	GAGCACAATGGAAACCGACCCCAACACACACAGCAACGCTCTCTCAACAATACTAATACA	1116						

RESULT 2

US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164

US-08-760-615-3
Query Match 74.0%; Score 1514.8; DB 3; Length 2164;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATAGTCTCATTTTAAATCCAAAGTATATAAAATCTCCCT 60
Db 119 ATGAAGACCATATTTCTTCTTATAGTCTTATCTATCTTAAATTCAGGACACAAAAATCTCCC 178
Qy 61 GTTTTAGAATGTAGTAAACAGCCAACTCAAGATGTAGATTGATGTGCTCCGGAACC 120
Db 179 ATTTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAC 238
Qy 121 CTCCAAAAGACAGAAGATGTTTCATCTGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 239 CTCGAAGACAGAAGAGCTCACTGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 298
Qy 181 GATTCCCTTTGGAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 299 GATTCCCTTTGGAGCATCAAGCGATGGCTTTTCAGGACAGGTGTTACCTCCCAAGAA 358
Qy 241 GTTAGATACGGAAGGAGAGAGCCAAACATGTTACATATAAGTGTAAAGACCCCT 300

Db 359 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAGTGTAAACGATCCC 418
Qy 301 TCTGAAAAATCCTTGTCTGTGATCCTCCAGTAAATATCCCGATTAACCTTAAATGTAAA 360
Db 419 TCTGAAAAATCCTTGTCTGTGATCCTCCAGTAAATATCCCGATTAACCTTAAATGTAAA 478
Qy 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTTCATGACAGGGGATTCGCTCATTTGTTGG 420
Db 479 ACTATCCATCATATTTCAAGGTCAAAACCTTCATGACAGGGGATTCGCTCATTTGTTGG 538
Qy 421 GGGGCATTTTCTTGTATGATCGGTTTGCCTCTCAACAATGTACCGAGGCAAGTCTTTC 480
Db 539 GGAGCATTTTCTTGTATGATCGGTTTGCCTCTCAACAATGTACCGAGGCAAGTCTTTC 598
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCAAGAAATGATTTTTCT 540
Db 599 ACTGAAGGGAACATAGCAGCTATGATTTGTTCAATAGACAGTTCAAGAAATGATTTTTCT 658
Qy 541 AGGCAAGGACAAAGGTTATCGTCATGAACTTGGACCTCCACCAATTAATATTGGACAGC 600
Db 659 CGGCAAGGACAAAGGTTATCGGTTTGGCATCTCTCCAGAAATACAACTCC 660
Qy 601 AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
Db 719 AGTAAACGGAACGCAAAACGATGACACTGGATGTTTGGCGCTCTTCAAGAAATACAA 778
Qy 661 ACAAACAAATCAAAACATGCTCTTAAACCTCCATCTTAAACCTCCATCCCTGCCACAGTAACTCCG 720
Db 779 ACAAAGAACCAAAACATGCTCTTAAACCTCCATCTTAAACCTCCATCCCTGCCACAGTAACTCCG 838
Qy 721 AGCATTCATCTACAATATCTCAAAATTAATATCTGCTAAATCTGGAACCTATGAACCCAACT 780
Db 839 GAGATCAAACTCAACAGCACCCCAACTGATGTCACCAACTCAATACACAGGACCAAGC 898
Qy 781 AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 899 AGTATGATGAGGACCTCGCAACATCGGCTCAGGTCGGGAGAACAGGAAACCCACACA 958
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAAATATTTGTCCACTCTTCACTA 900
Db 959 ACTTCTGATGGGTCAACAGCAAGGGCTTTTCATCAACAAATGCCACCTCCCTCACCA 1018
Qy 901 CATCCAAAGCCTCAACACATGAGCAAAAACGATGAGAACTCTTCCGACATCTGTGTAAT 960
Db 1019 CAACCAAGCAGCCACAGCAAGGAGGAAACAAACAAACCAATTCCTCCAGATGCTGTGACT 1078
Qy 961 GAGCACAATGGAACCGACCCCAACAAACAAACAGCGCTCCTCAACAATACTAATACA 1020
Db 1079 GAATAGACAAAATAACACAACTGCAACACCGTCCATGCCCTCTCATACACTACCA 1138
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCCGCAAC 1080
Db 1139 ATCTCTACTAACACACTCCAAACACAACTTCAGACTCTCTCTGACCACTTACAAAC 1198
Qy 1081 ATCAACCAATATGATACAAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
Db 1199 ACCACCAATGACACACAGAGCAAACTCACTGAAATGAGCAAAACAGTGCCTCCCTCG 1258
Qy 1141 AACCAACTCTAGATCCAAAGAAATCCCAACACAGCAAGACAGCAACCAAGCAGCAAC 1200
Db 1259 ATAACAACCTCTGCTCCAAAGGAAATCCCAACAGCAAGCAAGAGCAGCAGCAAGAAAA 1318
Qy 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAAAATTTCTTCCGGAT 1260
Db 1319 GGCCCCGCAACAGGACCAACACAGCAAAATGAGCAATTCACAGTCTCTCCCTCCAC 1378
Qy 1261 TCTAGTCCGCAACCCGCTCTTATATATATTTTGAAGAAAGAAACGATCGATCTCTGAGG 1320
Db 1379 CCCAGCTCGACTGCAACATCTTGTATATTTTCAAGAAAGCGAAGTATCTCTGAGG 1438
Qy 1321 GAAGCGCATGTTTCCCTTTCTGATGGGTTAATTAATGCTCCAAATGATTTTGAACCA 1380
Db 1439 GAAGCGCATGTTTCCCTTTCTGATGGGTTAATTAATGCTCCAAATGATTTTGAACCA 1498

```
QY 1381 GTTCAAATACAAACAAATCTTTGATGAATCTCTAGTCTCTAGTCTCTGGTGCCTCGGCTGAGGAA 1440
|||
Db 1499 GTTCAAATACAAACAAATCTTTGATGAATCTCTAGTCTCTAGTCTCTGGTGCCTCGGCTGAGGAA 1558
|||
QY 1441 GATCAACATGCTCCCTCCCAATATAGTTAGTTAACTTTATCTTTATCTTTTCCCTAAATATAATGAG 1500
|||
Db 1559 GATCAACATGCTCCCTCCCAATATAGTTAGTTAACTTTATCTTTATCTTTTCCCTAAATATAATGAG 1618
|||
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGTGATGAGAGTTAAGAAATTTGGAGC 1560
|||
Db 1619 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGTGATGAGAGTTAAGAAATTTGGAGC 1678
|||
QY 1561 GTTCAAGAGATGACTCTGGCCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
|||
Db 1679 GTTCAAGAGATGACTCTGGCCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1738
|||
QY 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAATTTGGCTCTGAGGTTGAGG 1680
|||
Db 1739 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAATTTGGCTCTGAGGTTGAGG 1798
|||
QY 1681 CGTCTAGCAATCAAACTGCCAAATCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1740
|||
Db 1799 CGTCTAGCAATCAAACTGCCAAATCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1858
|||
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCACAAGATGGGGAGG 1800
|||
Db 1859 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCACAAGATGGGGAGG 1918
|||
QY 1801 ACATGCAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTGTCCAAAATATT 1860
|||
Db 1919 ACATGCAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTGTCCAAAATATT 1978
|||
QY 1861 TCAGAGCAATTTGACCAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 1920
|||
Db 1979 TCAGAGCAATTTGACCAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 2038
|||
QY 1921 GGTGTAATGTTGACATCGGACTGGGGTGTCTTACTAACTTGGGCAATTTTGTCTACTA 1980
|||
Db 2039 GGTGTAATGTTGACATCGGACTGGGGTGTCTTACTAACTTGGGCAATTTTGTCTACTA 2098
|||
QY 1981 TTATCCATAGTCTTGAATGCTCTATCTGATTTGCTGATTTTCTGATTTTCTAAATATC 2040
|||
Db 2099 TTATCCATAGTCTTGAATGCTCTATCTGATTTTGTGATTTTGTGATTTTCTAAATATC 2158
|||
QY 2041 GGATAA 2046
|||
Db 2159 GGATAA 2164
|||
```

RESULT 3

```
US-09-336-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
```

```
; ORGANISM: Marburg Virus
US-09-336-910A-1
```

```
Query Match 74.0%; Score 1514.8; DB 4; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTTCTGATGATCTCTATTTTAAATCCAAAGTATAAAACTCTCCCT 60
Db 5940 ATGAAGACCATATATTTTCTGATGATCTCTATTTTAAATCCAAAGTATAAAACTCTCCCT 5999
|||
QY 61 GTTTTGAATTTCTAGTAAACAGCCCACTCAAGATCTAGATTCAGTCTGCTCCGGAACC 120
|||
Db 6000 ATTTTGAAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059
|||
QY 121 CTCCAAAGACAGAAAGATGTTTCACTCTGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
|||
Db 6060 CTCCAAAGACAGAAAGATGTTTCACTCTGATGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 6119
|||
QY 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
|||
Db 6120 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 6179
|||
QY 241 GTTGATGATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
|||
Db 6180 GTTGATGATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6239
|||
QY 301 TCTGGAATAATCTTGTCTGCTGGATCTCCCAAGTAAATATCCCGGATTTACCTTAAATGTAAA 360
|||
Db 6240 TCTGGAATAATCTTGTCTGCTGGATCTCCCAAGTAAATATCCCGGATTTACCTTAAATGTAAA 6299
|||
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTTCATGCAAGGGGATTTGCCCTCCATTTGTGG 420
|||
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTTCATGCAAGGGGATTTGCCCTCCATTTGTGG 6359
|||
QY 421 GGGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTTACCGAGGCAAGGCTCTTC 480
|||
Db 6360 GGGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTTACCGAGGCAAGGCTCTTC 6419
|||
QY 481 ACTCAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAAGAAATGATTTTCT 540
|||
Db 6420 ACTCAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAAGAAATGATTTTCT 6479
|||
QY 541 AGGCAAGGACAAAGTTATCGTCAATGAACCTTACCTCCCAATATAATATTTGGAACAAG 600
|||
Db 6480 CGGCAAGGACAAAGGTTATCGTCAATGAACCTTACCTCCCAATATAATATTTGGAACAAG 6539
|||
QY 601 AGCAATGAACGAGAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
|||
Db 6540 AGTAACGGAACGAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 6599
|||
QY 661 ACAAAACATCAAAACATGCTCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 720
|||
Db 6600 ACAAGGAACCAAAACATGCTCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 6659
|||
QY 721 AGCATTTCATCTACAAATFACTCAAAATTAATACTGCTAAATCTGGAACCTATGAACCAAGT 780
|||
Db 6660 GAGATCAAACTCAAGCACCCTCAAGTGTCCCAACCAAACTCAATACCAAGCACCCTCAAG 6719
|||
QY 781 AGCAGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAAACAGGGGCCCCACACA 840
|||
Db 6720 AGTGAATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAAACAGGAAACCCCAACA 6779
|||
QY 841 ACTCTTAATGATGACTGAAACGAAACAACTGCTCAACAATATTGTCACCTCTCTTCACTA 900
|||
Db 6780 ACTTCTGATGGGTCAACCAAGAGGGCTTTTATCAACATGCCCACCTCCCTCACCACCA 6839
|||
QY 901 CATCCAGCACCTCACAACATGAGCAAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
|||
Db 6840 CAACCAAGCACGCCACAGCAGGAGGAGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 6899
|||
QY 961 GAGCACATGGAACCGGACCCCAACACACACAGCAGGCTCTCTCAACAATACTAATACA 1020
|||
```

Db 6900 GAACCTAGACAAAATAACACAACTGTCACACCGTCTCCCTCCCTCATAACACACTACCACA 6959
QY 1021 ACTCCACCTATACACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCTCCAAACCGGCAAC 1080
Db 6960 ATCTCTACTAACACACCTCCAAACAACTTTCAGCACTCTCTCTGCAACCAATTACAAAC 7019
QY 1081 ATACCAATAATGATACACAACTGAACTAGCAGAAAGCGAAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACACACACAGACACATCACTGAAATGAGCAACCACTGCTCCCTCG 7079
QY 1141 AACCAACTCTAGATCCAAAGAAAATCCACACAGGACAGACACCAACAGACACCAAC 1200
Db 7080 ATAACAACCTCTCCAAAGGAAATCCACACAGCAAGAGACAGCAGCAGCAAGAAA 7139
QY 1201 AACATCATGACGACATCAGATATACAGAGCAACACCCCAACAATCTTCTCCGAT 1260
Db 7140 GGCCCGCGCACAGCGGACCAACACGCAATGAGCATTTTCCAGCTCTCCGCCAC 7199
QY 1261 TCTAGTCGACACACCGCCCTCTATATACCTTGAAGAAAGACGATCGATCCTCTGGAGG 1320
Db 7200 CCCAGCTCGATGACAAACATCTTGTATATTTGAGAAAGCGAAGTATCTCTGGAGG 7259
QY 1321 GAAGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 1380
Db 7260 GAAGCGACATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 7319
QY 1381 GTTCCAAATACAAAACAATCTTGTAGTAATCTCTAGTTCTGGTGGCTCGGCTGAGGAA 1440
Db 7320 GTTCCAAATACAAAACAATCTTGTAGTAATCTCTAGTTCTGGTGGCTCGGCTGAGGAA 7379
QY 1441 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCTTAAATAAATGAG 1500
Db 7380 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCTTAAATAAATGAG 7439
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTCAGACAGAGTTAAGAAATTTGGAGC 1560
Db 7440 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTCAGACAGAGTTAAGAAATTTGGAGC 7499
QY 1561 GTTCAGGAGATGACTCGGCGAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAAT 1620
Db 7500 GTTCAGGAGATGACTCGGCGAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAAT 7559
QY 1621 GAAGGACTTTACCTGCTGTTTAAATAAATCAAAACAAATTTGGTCTGCAGGTTGAGG 1680
Db 7560 GAAGGACTTTACCTGCTGTTTAAATAAATCAAAACAAATTTGGTCTGCAGGTTGAGG 7619
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Db 7620 CGTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 7679
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTTGCTTCTTACTCACAAGATGGGAGGA 1800
Db 7680 AGAATCTCTCTTAATCAATAGACATGCTATTTGCTTCTTACTCACAAGATGGGAGGA 7739
QY 1801 ACATGCAAGTCTTGGACCTGATTTGCTATCGGATAGAGACTTGTCCAAAATATT 1860
Db 7740 ACATGCAAGTCTTGGACCTGATTTGCTATCGGATAGAGACTTGTCCAAAATATT 7799
QY 1861 TCAGAGCAATTTGACCAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTCG 1920
Db 7800 TCAGAGCAATTTGACCAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTCG 7859
QY 1921 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTTGTCTACTA 1980
Db 7860 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTTGTCTACTA 7919
QY 1981 TTATCCATAGTCTTCTGATGCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 2040
Db 7920 TTATCCATAGTCTTCTGATGCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 7979
QY 2041 GGATAA 2046
Db 7980 GGATAA 7985

RESULT 4
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172
; OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1

Query Match 6.5%; Score 134; DB 3; Length 2172;
Best Local Similarity 60.9%; Pred. No. 1.2e-30;
Matches 218; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1532 ATTCTGATGACAGCTTAAAGAAATTTGGAGCGTTTCAAGGAGATGACTGGCCGCGAGGGCTCA 1591
Db 1670 AATGCAACCCCTAATTTACATTAATCTAGGATGAAGGTGCTGCAATCGGACTGG 1729
QY 1592 GTTGGATACCGTTTTTGGCCCTCGAAATTTGAAGGACTTTACACTGCTGTTTTTAATTAATA 1651
Db 1730 CCTGGATACCATATTTTCGGGCCAGCAGCCGAGGGAAATTTACATAGAGGGGCTAATGCACA 1789
QY 1652 ATCAAAACAATTTGGTCTGCAGGTTGAGGCGTCTAGCCAACTCAAACTGCCAAATCCTTGG 1711
Db 1790 ATCAAGATGGTTAATCTGTGGTTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTC 1849
QY 1712 AACTCTTATTGAGAGTCACAACTGAGGAAAGAAACATTTCTCTTAAATCAATAGACATGCTA 1771
Db 1850 AACTGTTCTTGAGAGCCCAACTGAGTACGCACTTTTCAATCTTCAACCGTAAGGCA 1909

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verry, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5163 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-4

Query Match 2.6%; Score 54; DB 3; Length 5163;
Best Local Similarity 44.6%; Pred. No. 2e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCACACAACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTGTCCACT 891
Db 544 CCATACACTAAATGTGTTGGAGTGAACACACACACACACACAACTACTACTACT 603

Qy 892 CTTCACTACATCCAGACCTCAACATGAGCAAAACAGTACGATCCTTCCGACAT 951
Db 604 ACTACTACGACAAACACAAACACGACACAACTACTACTACTACTACTACTACT 663

Qy 952 GCTGTAAGTGAACAAATGGAACCGACCCACACACACACACAACTCTCTCAACAAT 1011
Db 664 ACTACTACGACAAACACAAACACAAACACAAACACAAACAAACAACA 723

Qy 1012 ACTAATACAACTCCCACTTAACTCTCAAGTACAACTCTCAAGTACAACTCTCTTCCCTCCA 1071
Db 724 ACAACAAACGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 783

Qy 1072 ACCCGCAATCACAATAATGATACAAAGTGAATAGCAGAAAGGGAACAACCAAT 1131
Db 784 ACAACAACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACT 843

Qy 1132 GCTCAGTTGAACAACTCTAGATCCACAGAAATCCACACAGGACAGACCAAC 1191
Db 844 ACACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 903

Qy 1192 AGCAACCAACATCATCATGACGACATCAGATATAACAGCAAAACCCCAAAATTTCT 1251
Db 904 ACACTACCAAGAAACCAACAAACAACTACTACTACTACTACTACTACTACTACT 963

Qy 1252 TCTCCGGATTCTAGTCCGACAAACCGCCCTCTATATATCTTTAGAAAGAAACGATCGA 1309
Db 964 ACTAACAACCAACCAACCAACCAACCAACCAACCACTACTACTACTACTACTACT 1021

RESULT 8
US-08-928-361B-4
; Sequence 4, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5

; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5163
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-4

Query Match 2.6%; Score 54; DB 4; Length 5163;
Best Local Similarity 44.6%; Pred. No. 2e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCACACAACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTGTCCACT 891
Db 544 CCATACACTAAATGTGTTGGAGTGAACACACACACACACAACTACTACTACTACTACT 603

Qy 892 CTTCACTACATCCAGACCTCAACATGAGCAAAACAGTACGATCCTTCCGACAT 951
Db 604 ACTACTACGACAAACACAAACACGACACAACTACTACTACTACTACTACTACTACT 663

Qy 952 GCTGTAAGTGAACAAATGGAACCGACCCACACACACACAACTCTCTCAACAAT 1011
Db 664 ACTACTACGACAAACACAAACACAAACACAAACACAAACAAACAACA 723

Qy 1012 ACTAATACAACTCCCACTTAACTCTCAAGTACAACTCTCAAGTACAACTCTCTTCCCTCCA 1071
Db 724 ACAACAAACGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 783

Qy 1072 ACCCGCAATCACAATAATGATACAAAGTGAATAGCAGAAAGGGAACAACCAAT 1131
Db 784 ACAACAACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACT 843

Qy 1132 GCTCAGTTGAACAACTCTAGATCCACAGAAATCCACACAGGACAGACCAAC 1191
Db 844 ACACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 903

Qy 1192 AGCAACCAACATCATCATGACGACATCAGATATAACAGCAAAACCCCAAAATTTCT 1251
Db 904 ACACTACCAAGAAACCAACAAACAACTACTACTACTACTACTACTACTACTACTACT 963

Qy 1252 TCTCCGGATTCTAGTCCGACAAACCGCCCTCTATATATCTTTAGAAAGAAACGATCGA 1309
Db 964 ACTAACAACCAACCAACCAACCAACCAACCAACCACTACTACTACTACTACTACTACT 1021

RESULT 9
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-4 (HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; CURRENT FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; EARLIER FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 15

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5318
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-08-700-651-2

Query Match      2.6%; Score 54; DB 3; Length 5318;
Best Local Similarity 44.6%; Pred. No. 2.1e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCGACAACTCTTAATGTAGTCACTGAAAGAAACAATCGTCAACAATATTGTCCACT 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 CCATACACTAAATGTGTGGAGTGAACACACAAACAACAACAACAACAACAACAACA 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 892 CTTTCACTACATCCAAAGCACCTCAACAATGAGCAAAACAGTACGAATCTTCCGACAT 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 ACTACTAGCAACAACAACAACAACAAGACACAACAACAACAACAACAACAACAACA 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 952 GCTGTAACTGAGCAACAATGGAAACCGACCAACAACAACAACAACAACAACAACA 1011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 ACTACTAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 724
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1012 ACTAATAAATCCACCTTAACTACTCTCAAGTACAACCTCAGTACTCTTCCCTCCA 1071
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 ACAACAACAAGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 784
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1072 ACCCGCACTCACAATAATGATACACAACGTGAACCTAGCAGAAAGCGAACAACCAAT 1131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 ACAACAATACAAACCAACTACAAACCAACAACAACAACAACAACAACAACAACA 844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1132 GTCAGTTGAAACACAACCTCTAGATCCAAACAGAAAATCCCAACAGGACAGACCAAC 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 ACAACTACAAACCAACTACAAACCAACAACAACAACAACAACAACAACAACAAC 904
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1192 AGCAACAACAATCATCATGACGACATCAGATATAACAAGCAAAACCCCAACAATTTCT 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ACAACTACCAAGAAACCAACAACAACAACAACAACAACAACAACAACAACAACA 964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1252 TCTCCGGATTCTAGTCCGACACCCGCCCTCTATATACTTTAGAAAGAAACGATCGA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 ACTAACCAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1022
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS.
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 60/026,062
```

```
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verry, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-3

Query Match      2.6%; Score 54; DB 3; Length 5318;
Best Local Similarity 44.6%; Pred. No. 2.1e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCGACAACTCTTAATGTAGTCACTGAAAGAAACAATCGTCAACAATATTGTCCACT 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 CCATACACTAAATGTGTGGAGTGAACACACAACAACAACAACAACAACAACAACA 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 892 CTTTCACTACATCCAAAGCACCTCAACAATGAGCAAAACAGTACGAATCTTCCGACAT 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 ACTACTAGCAACAACAACAACAACAAGACACAACAACAACAACAACAACAACAACA 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 952 GCTGTAACTGAGCAACAATGGAAACCGACCAACAACAACAACAACAACAACAACA 1011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 ACTACTAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1012 ACTAATAAATCCACCTTAACTACTCTCAAGTACAACCTCAGTACTCTTCCCTCCA 1071
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 ACAACAACAAGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1072 ACCCGCACTCACAATAATGATACACAACGTGAACCTAGCAGAAAGCGAACAACCAAT 1131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 ACAACAATACAAACCAACTACAAACCAACAACAACAACAACAACAACAACAACA 843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1132 GCTCAGTTGAAACAACAATCTAGATCCAAACAGAAAATCCCAACAGGACAGACCAAC 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 ACAACTACAAACCAACTACAAACCAACAACAACAACAACAACAACAACAACAACC 903
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1192 AGCAACAACAATCATCATGACGACATCAGATATAACAAGCAAAACCCCAACAATTTCT 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 904 ACAACTACCAAGAAACCAACAACAACAACAACAACAACAACAACAACAACAACA 963
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1252 TCTCCGGATTCTAGTCCGACACCCGCCCTCTATATACTTTAGAAAGAAACGATCGA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 964 ACTAACCAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-588-995A-3
; Sequence 3, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
```


; PRIOR APPLICATION NUMBER: 08/700,651
 ; PRIOR FILING DATE: 1996-08-14
 ; PRIOR APPLICATION NUMBER: 08/415,751
 ; PRIOR FILING DATE: 1995-04-03
 ; NUMBER OF SEQ ID NOS: 115
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 5318
 ; TYPE: DNA
 ; ORGANISM: *Cryptosporidium parvum*
 US-09-588-995A-3

Query Match 2.6%; Score 54; DB 4; Length 5318;
 Best Local Similarity 44.6%; Pred. No. 2.1e-05;
 Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

QY	832	CCCCACAACTCTTAAATGTAGTCACCTGAACAGAGAAACAATCGTCAACAATATTGTGCACT	891
DB	544	CCATACACTAAATGTGTGGAGTGAACACACACAAACACAACTACTACTACTACTACT	603
QY	892	CCTTCACTCATCCAAAGCACCTCAACAACATGAGCAAAACAGTAGTCAATCTCTTCCCGACAT	951
DB	604	ACTTACTACGACAAACAACAACAAACGACAAACAACTACTACTACTACTACTACTACT	663
QY	952	GCTGTAACTGAGCACAATGGAACCGACCCACAAACAACAACAGCAACGCTCTTCAACAAT	1011
DB	664	ACTTACTACGACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA	723
QY	1012	ACTAATACAACTCCCACTTATACACTCTCAAGTACAACTCTAGTACTCTTTCCTCCCA	1071
DB	724	ACAACAACAACGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT	783
QY	1072	ACCGCAACATCAACCAATATGATACAAACAGTGAACTAGCAGAAAGCGCAACAAACCAAT	1131
DB	784	ACAACAACACTACACCACAACATCAACCCACAACAACACTACAACCAACAACCAACCA	843
QY	1132	GCTCAGTTGAACCAACTCTAGATCCAAAGAAAATCCCAACAGGACAGACACCAAC	1191
DB	844	ACAACACTAACACCACAACATCAACCCACAACAACACTACAACCAACAACCAACCAAC	903
QY	1192	AGCACACCAACATCATCATGACGACATCAGATATACAGCAACACCCCAACAATCT	1251
DB	904	ACAACACTAACAGAAACCAACAACAACAACTACTACAACAACAACAACAACAACAACA	963
QY	1252	TCTCCGGATTCTTAGTCGCAACCCGCCCTCTATATCTTTAGAAAGAACGATCGA	1309
DB	964	ACTACAACCAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA	1021

RESULT 12
 US-09-270-767-2848/c
 ; Sequence 2848, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2848
 ; LENGTH: 631
 ; TYPE: DNA
 ; ORGANISM: *Drosophila melanogaster*
 ; FEATURE:
 ; OTHER INFORMATION: n means any nucleotide
 US-09-270-767-2848

Query Match 2.4%; Score 48.8; DB 4; Length 631;
 Best Local Similarity 51.1%; Pred. No. 0.0002;
 Matches 139; Conservative 0; Mismatches 132; Indels 1; Gaps 1;

;/ TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;/ TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;/ TITLE OF INVENTION: SPECIES INFECTIONS
;/ NUMBER OF SEQUENCES: 30
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: PETERS, VERNY, JONES & BIKSA
;/ STREET: 385 Sherman Avenue, Suite 6
;/ CITY: Palo Alto
;/ STATE: CA
;/ COUNTRY: USA
;/ ZIP: 94306-1840
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: PC-DOS/MS-DOS
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/928,361B
;/ FILING DATE: 12-SEP-1997
;/ CLASSIFICATION:
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 60/026,062
;/ FILING DATE: 13-SEP-1996
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Verny, Hana
;/ REGISTRATION NUMBER: 30,518
;/ REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 650-324-1677
;/ TELEFAX: 650-324-1678
;/ INFORMATION FOR SEQ ID NO: 2:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 5511 base pairs
;/ TYPE: nucleic acid
;/ STRANDEDNESS: double
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: DNA (genomic)
;/ US-08-928-361B-2

Query Match 2.3%; Score 48; DB 3; Length 5511;
Best Local Similarity 44.2%; Pred. No. 0.0016;
Matches 198; Conservative 0; Mismatches 250; Indels 0; Gaps 0;
Qy 838 ACAACTCTTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTCA 897
Db 988 ACAACAATCTACTACACTACTACCTACTAGGACAAACAATCTACTACTACTACTACTAGCA 1047
Qy 898 CTACATCCAGCACCTCACAATGAGCAAAACAGTACGAATCTTCCGACATGCTGTA 957
Db 1048 ACAACAACAATCTACTACTACTACTACTACCAACCACTACCAACCAACCACTACTACTA 1107
Qy 958 ACTGAGCAATGGAACCGACCAACCAACCAACCAACCAACCAACCAACCAACCACTAAT 1017
Db 1108 ACCAACAACAATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1167
Qy 1018 ACAACTCCACCTATAACACTCTCAAGTACAACTCTCAAGTACAACTCTCTTCCCTCCACCGC 1077
Db 1168 ACA 1227
Qy 1078 AACATCAACAATATGATACACAACGTTGAACTAGCAGAAAGCGAACAACCAATGCTCAG 1137
Db 1228 ACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1287
Qy 1138 TTGAACAACAATCTAGATCCAAAGAAATCCCAACGAGCAAGACCAACAACAACAACAACA 1197
Db 1288 ACAACAACAACAATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1347
Qy 1198 ACCAACAATCATGACGACATGATATAACAAGCAAAACCCCAAAAATTTCTTCGG 1257
Db 1348 ACTACAATCTACAGAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1407
Qy 1258 GATTCTAGTCCGAAACCGCCCTCCTA 1285
Db 1408 ACTACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1435

Search completed: August 1, 2005, 10:31:17
Job time : 332.71 secs

Db 1408 ACTACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1435
RESULT 15
; Sequence 2, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5511
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
; US-09-588-995A-2

Query Match 2.3%; Score 48; DB 4; Length 5511;
Best Local Similarity 44.2%; Pred. No. 0.0016;
Matches 198; Conservative 0; Mismatches 250; Indels 0; Gaps 0;
Qy 838 ACAACTCTTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTCA 897
Db 988 ACAACAATCTACTACACTACTACCTACTAGGACAAACAATCTACTACTACTACTACTAGCA 1047
Qy 898 CTACATCCAGCACCTCACAATGAGCAAAACAGTACGAATCTTCCGACATGCTGTA 957
Db 1048 ACAACAACAATCTACTACTACTACTACTACCAACCACTACCAACCAACCACTACTACTA 1107
Qy 958 ACTGAGCAATGGAACCGACCAACCAACCAACCAACCAACCAACCAACCAACCACTAAT 1017
Db 1108 ACCAACAACAATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1167
Qy 1018 ACAACTCCACCTATAACACTCTCAAGTACAACTCTCAAGTACAACTCTCTTCCCTCCACCGC 1077
Db 1168 ACA 1227
Qy 1078 AACATCAACAATATGATACACAACGTTGAACTAGCAGAAAGCGAACAACCAATGCTCAG 1137
Db 1228 ACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1287
Qy 1138 TTGAACAACAATCTAGATCCAAAGAAATCCCAACGAGCAAGACCAACAACAACAACAACA 1197
Db 1288 ACAACAACAACAATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1347
Qy 1198 ACCAACAATCATGACGACATGATATAACAAGCAAAACCCCAAAAATTTCTTCGG 1257
Db 1348 ACTACAATCTACAGAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1407
Qy 1258 GATTCTAGTCCGAAACCGCCCTCCTA 1285
Db 1408 ACTACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1435

Search completed: August 1, 2005, 10:31:17
Job time : 332.71 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1285.84 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatattttct.....ttactaatatcggataa 2046

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	15 US-10-066-506A-7	Sequence 7, Appl
2	1847.6	90.3	2046	15 US-10-066-506A-13	Sequence 13, Appl
3	1514.8	74.0	7778	20 US-10-491-121-30	Sequence 30, Appl
4	1514.8	74.0	11460	15 US-10-267-322-1	Sequence 1, Appl
5	1485.6	72.6	2051	15 US-10-066-506A-9	Sequence 9, Appl
6	1473.2	72.0	2046	21 US-10-811-353-2	Sequence 2, Appl
7	1473.2	72.0	19112	17 US-10-353-856-27	Sequence 27, Appl

8	1418.8	69.3	7005	20 US-10-491-121-31	Sequence 31, Appl
9	1418.8	69.3	8256	20 US-10-491-121-32	Sequence 32, Appl
10	1319.6	64.5	2046	15 US-10-066-506A-5	Sequence 5, Appl
11	834	40.8	6902	20 US-10-491-121-42	Sequence 42, Appl
12	780	38.1	1841	15 US-10-066-506A-3	Sequence 3, Appl
13	746.6	36.5	2252	15 US-10-066-506A-1	Sequence 1, Appl
14	149	7.3	18890	17 US-10-353-856-9	Sequence 9, Appl
15	148.2	7.2	6885	20 US-10-491-121-18	Sequence 18, Appl
16	148.2	7.2	6889	20 US-10-491-121-19	Sequence 19, Appl
17	148.2	7.2	7002	20 US-10-491-121-16	Sequence 16, Appl
18	148.2	7.2	7003	22 US-10-860-878-1	Sequence 1, Appl
19	148.2	7.2	7023	20 US-10-491-121-21	Sequence 21, Appl
20	148.2	7.2	7036	20 US-10-491-121-17	Sequence 17, Appl
21	148.2	7.2	8146	20 US-10-491-121-20	Sequence 20, Appl
22	145.8	7.1	2360	17 US-10-397-635-12	Sequence 12, Appl
23	145.8	7.1	6913	20 US-10-491-121-11	Sequence 11, Appl
24	145.8	7.1	8131	20 US-10-491-121-12	Sequence 12, Appl
25	134.4	6.6	6940	20 US-10-491-121-15	Sequence 15, Appl
26	134.4	6.6	7073	22 US-10-860-878-2	Sequence 2, Appl
27	134.4	6.6	7082	20 US-10-491-121-13	Sequence 13, Appl
28	134.4	6.6	7087	20 US-10-491-121-14	Sequence 14, Appl
29	134.4	6.6	8221	20 US-10-491-121-27	Sequence 27, Appl
30	134.4	6.6	8338	20 US-10-491-121-26	Sequence 26, Appl
31	134.4	6.6	10783	20 US-10-491-121-25	Sequence 25, Appl
32	134	6.5	2039	15 US-10-066-506A-11	Sequence 11, Appl
33	134	6.5	2298	9 US-09-337-946A-1	Sequence 1, Appl
34	134	6.5	2298	17 US-10-384-976-1	Sequence 1, Appl
35	134	6.5	2298	18 US-10-226-795-1	Sequence 1, Appl
36	134	6.5	2298	22 US-10-696-633-1	Sequence 1, Appl
37	134	6.5	6467	20 US-10-491-121-10	Sequence 10, Appl
38	134	6.5	6561	20 US-10-491-121-4	Sequence 4, Appl
39	134	6.5	6624	20 US-10-491-121-3	Sequence 3, Appl
40	134	6.5	6914	20 US-10-491-121-9	Sequence 9, Appl
41	134	6.5	7154	20 US-10-491-121-1	Sequence 1, Appl
42	134	6.5	7188	20 US-10-491-121-2	Sequence 2, Appl
43	134	6.5	7272	22 US-10-860-878-4	Sequence 4, Appl
44	134	6.5	7285	22 US-10-860-878-3	Sequence 3, Appl
45	134	6.5	8199	20 US-10-491-121-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-10-066-506A-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION: Case C.
; APPLICANT: Grogan, Michael C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein
; OTHER INFORMATION: Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-7

Query Match 100.0%; Score 2046; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 60

[illegible]

RESULT 2
US-10-066-506A-13
; Sequence 13, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.

APPLICANT: Hevey, Michael C.
INVENTOR: Schmaljohn, Alan, L.
TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
FILE REFERENCE: 003/243/SAP
CURRENT APPLICATION NUMBER: US/10/066,506A
CURRENT FILING DATE: 2002-01-31
PRIOR FILING DATE: 2002-01-31
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 13
LENGTH: 2046
TYPE: DNA
ORGANISM: Marburg virus strain Raven
FEATURE:
OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg virus Glycoprotein 2
OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-13

Query Match 90.3%; Score 1847.6; DB 15; Length 2046;

Best Local Similarity 93.9%; Pred. No. 0;
Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy	1	ATGAAGACCATATATTTCTGATGCTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT	60
Db	1	ATGAAGACCATATATTTCTGATGCTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT	60
Qy	61	GTTTTAGAAATGTGTAGTAAACAGCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC	120
Db	61	GTTTTAGAAATGTGTAGTAAACAGCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC	120
Qy	121	CTCCAAAGACAGAGATGTTCTCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT	180
Db	121	CTCCAAAGACAGAGATGTTCTCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT	180
Qy	181	GATTCCTCTTGGAGGATCTAAACGATGGGCTTTCAGGACAGTGTTCCTCCCAAGAAC	240
Db	181	GATTCCTCTTGGAGGATCTAAACGATGGGCTTTCAGGACAGTGTTCCTCCCAAGAAC	240
Qy	241	GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT	300
Db	241	GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT	300
Qy	301	TCTGGAAATCTTGCTGCTGGATCTCCAGTAAATATCCGCAATACCTTAATGTAAA	360
Db	301	TCTGGAAATCTTGCTGCTGGATCTCCAGTAAATATCCGCAATACCTTAATGTAAA	360
Qy	361	ACTGTTCAATATTCAGGTCAAACCTCATGCACAGGGGATTCCTCCATTTGTGG	420
Db	361	ACTGTTCAATATTCAGGTCAAACCTCATGCACAGGGGATTCCTCCATTTGTGG	420
Qy	421	GGGGCATTTTCTTGATGATCGGTTGCTCTACAAATGTATCCGAGGCAAGTCTTC	480
Db	421	GGGGCATTTTCTTGATGATCGGTTGCTCTACAAATGTATCCGAGGCAAGTCTTC	480
Qy	481	ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTTCAAGAAATGATTTTCT	540
Db	481	ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTTCAAGAAATGATTTTCT	540
Qy	541	AGGCAAGACAGGTTATCGTCACATGAACTTGACCTCCACCAATATATGGACAGC	600
Db	541	AGGCAAGACAGGTTATCGTCACATGAACTTGACCTCCACCAATATATGGACAGC	600
Qy	601	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC	660
Db	601	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC	660
Qy	661	ACAAACAAATCAACATGCCCTCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG	720
Db	661	ACAAACAAATCAACATGCCCTCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG	720
Qy	721	AGCATTCATCTACAAATCTCAATTAATCTGCTAAATCTGGAACTATGAACCCCAAGT	780
Db	721	AGCATTCATCTACAAATCTCAATTAATCTGCTAAATCTGGAACTATGAACCCCAAGT	780

Db	721	AGCATTCATCTACAAATCTCAATTAATCTGCTAAATCTGGAACTATGAACCCCAAGT	780
Qy	781	AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Db	781	AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Qy	841	ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTCACTA	900
Db	841	ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTCACTA	900
Qy	901	CATCCAGCACCTCAACAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT	960
Db	901	CATCCAGCACCTCAACAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT	960
Qy	961	GAGCACAATGGAACCGACCCCAACAAACAGTACGAATCTTCCCGACATGCTGTAACT	1020
Db	961	GAGCACAATGGAACCGACCCCAACAAACAGTACGAATCTTCCCGACATGCTGTAACT	1020
Qy	1021	ACTCCACCTTATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCCGCAAC	1080
Db	1021	ACTCCACCTTATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCCGCAAC	1080
Qy	1081	ATCACCATAATGATACACAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
Db	1081	ATCACCATAATGATACACAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
Qy	1141	AACACAACTCTAGATCCAAACAGAAAATCCCAACAGGACAAAGACACCAACAGCAACAC	1200
Db	1141	AACACAACTCTAGATCCAAACAGAAAATCCCAACAGGACAAAGACACCAACAGCAACAC	1200
Qy	1201	AACATCATGATGACGATATACAGCAATATACAGCAATATACAGCAATATCTTCTCCGAT	1260
Db	1201	AACATCATGATGACGATATACAGCAATATACAGCAATATACAGCAATATCTTCTCCGAT	1260
Qy	1261	TCTAGTCCGACAAACCCGCTCTCTATATATCTTGAAGAAAGAAACGATCGATCTCTGAGG	1320
Db	1261	TCTAGTCCGACAAACCCGCTCTCTATATATCTTGAAGAAAGAAACGATCGATCTCTGAGG	1320
Qy	1321	GAAGGACATGTTCTCTGATGGGTTTAAATGCTCCAAATGATTTGACCCCA	1380
Db	1321	GAAGGACATGTTCTCTGATGGGTTTAAATGCTCCAAATGATTTGACCCCA	1380
Qy	1381	GTTCCAATACAAACAAATCTTGTGATGAACTCTCTAGTTCTGGTCTCGGCTGAGAA	1440
Db	1381	GTTCCAATACAAACAAATCTTGTGATGAACTCTCTAGTTCTGGTCTCGGCTGAGAA	1440
Qy	1441	GATCAACATGCTCTCCCAATATATGATTTAACTTTTCTTATTTCTTAATAATAATAG	1500
Db	1441	GATCAACATGCTCTCCCAATATATGATTTAACTTTTCTTATTTCTTAATAATAATAG	1500
Qy	1501	AACATGCTCTCTGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATG	1560
Db	1501	AACATGCTCTCTGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATG	1560
Qy	1561	GTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1620
Db	1561	GTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1620
Qy	1621	GAAGGATCTTACATGCTGTTTAAATTAATAATAATAATAATAATAATAATAATAATA	1680
Db	1621	GAAGGATCTTACATGCTGTTTAAATTAATAATAATAATAATAATAATAATAATAATA	1680
Qy	1681	CGTCTAGCAATCAAACTGCCAAATCTCTTGAATCTTCTTGAAGTCAACAACTGAGAA	1740
Db	1681	CGTCTAGCAATCAAACTGCCAAATCTCTTGAATCTTCTTGAAGTCAACAACTGAGAA	1740
Qy	1741	AGAACATCTCTTAATCAATAGACATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1800
Db	1741	AGAACATCTCTTAATCAATAGACATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1800
Qy	1801	ACATGCAAGGCTTGGACCTGATTTGTCATCGGATAGAGAACTTCTTCAAAAATAT	1860
Db	1801	ACATGCAAGGCTTGGACCTGATTTGTCATCGGATAGAGAACTTCTTCAAAAATAT	1860

```
QY 1861 TCAGAGCAAAATTGACCAAAATTAAAAAGGACGAACAAAAAGAGGGGACTGGTTGGGGTCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGAGGAGAACTGGCTGGGGTCTA 1920
QY 1921 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTAACTTGGGGATTTTGGTACTA 1980
Db 1921 GGTGCAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGCATCTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTCTGATGCTCTATCCCTGTATTTGCTGTATCTTTTACTATAATATC 2040
Db 1981 TTATCTATAGCTGTTCTGATGCTCTGCTCTGTATCTGTCTGTATCTTCTCACTAAATACATT 2040
QY 2041 GGATAA 2046
Db 2041 GGATGA 2046

RESULT 3
US-10-491-121-30
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30

Query Match 74.0%; Score 1514.8; DB 20; Length 7778;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 2034 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAAATCAGGAGCAAAAATCTCCCC 2093
QY 61 GTTTTAGAAATGCTAGTAACGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAAACC 120
Db 2094 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGTATGCTCCGGAACT 2153
QY 121 CTCGAAAGACAGAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
Db 2154 CTCGAAAGACAGAAGACGTCATCTGATGGGATTTTCACTGAGTGGGCAAAAAGTTGCT 2213
QY 181 GATTCCTCTTTGGAAGCATATAAAGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 2214 GATTCCTCTTTGGAAGCATATAAAGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 2273
QY 241 GTTGTAGTATACGGAAGGAGAAAGCCAAAACATGTTTACAATATATAGTGTAAACAGACCT 300
Db 2274 GTTGTAGTATACGAGGGGGAGGAAGCCAAAACATGCTACATATAAGTGTAAAGGATCCC 2333
QY 301 TCTGAAAATCTTCTGCTGTGATCCTCCAGTAATATCCGGGATTAACCTTAATGTAAA 360
Db 2334 TCTGAAAATCTTCTGCTGTGATCCTCCAGTAATATCCGGGATTAACCTTAATGTAAA 2393
```

```
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATTTGCCCTCCATTTGTGG 420
Db 2394 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATTCGCCCTTCATTTATGG 2453
QY 421 GGGGCATTTTCTTGTATGATCGCGTTGCCCTCTACAACAATGTATACGAGGGAAGGTCCTTC 480
Db 2454 GGAGCATTTTCTTGTATGATCGCATTTGCCCTCCACAACAATGTATACGAGGGAAGGTCCTTC 2513
QY 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGAAATGATTTTTTCT 540
Db 2514 ACTGAAGGAACATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGAAATGATTTTTCTCG 2573
QY 541 AGGCAAGGACAAAGGTTATCGTCAATGAACCTTGACCTCCACCAATAAATATTGGACAAGC 600
Db 2574 CGGCAAGGACAAAGGTTATCGTCAATGAATCTGACTTCTACTAATAAATATTGGACAAGT 2633
QY 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTTTGGCATCTCCCAAGAATACAACTCC 660
Db 2634 AGTAAACGGAACGCAAGCAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATTTCT 2693
QY 661 ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
Db 2694 ACAAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCACTGCCACAGCCGCTCG 2753
QY 721 AGCATTCACCTCTACAAATATCTCAAAATTAATACTGCTAAATCTGGAACCTATGAACCCAAAGT 780
Db 2754 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAGGCCCAAGC 2813
QY 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACTGAGAGAACAGGGGGCCCCACACA 840
Db 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAAACGAGAAACCCCAACA 2873
QY 841 ACTCTTAATGTAGTCACTGGAACAGAAACAATCGTCAACAATATTGTCCACCTCTTCACTA 900
Db 2874 ACTTCTGATGGGTCCAAAGCAAGGGTTCATCAACAATGCCACCCCACTCCCTCACCA 2933
QY 901 CATCCAAGCACCTCACAACTATGACAAAACAGTAGTACGAATCTTCCCGACATGCTGTAACT 960
Db 2934 CAACCAAGCACGCCACAGCAAGGAGGAAACAACAACCAATTCCTCAAGATGCTGTGACT 2993
QY 961 GAGCAAAATGGAACCGACCCCAACAACAACAGCAAGCGTCTCTCAACAATATCTAATACA 1020
Db 2994 GAACTAGACAAAAATAACAACTGCAACCGTCCATGCCCTCCCTCATAAACACTACCA 3053
QY 1021 ACTCCCACTTATACATCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 3054 ATCTCTACTAACAAACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTAACAAC 3113
QY 1081 ATCAACCAATATGATACAAACGTCGAATAGCAGAAAGCGAAACCAACCAATGCTCAGTTG 1140
Db 3114 ACCCAATGACAAACACAGAGCACAATCACTGAAATGAGCAAAACCAAGTGCCTCCCTCG 3173
QY 1141 AACCAACTCTAGATCCAAAGAAAAATCCCAACAGAGCAAGAGACCAACAGCAACAC 1200
Db 3174 ATAACAACCTCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGAGCAAGCAAGCA 3233
QY 1201 AACATCATGAGCAGATCAGATATACAGAGCAAAACACCCCAACAATTTCTTCCGGAT 1260
Db 3234 GGCCCCGCAACAGGGCACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCTCCAC 3293
QY 1261 TCTAGTCCGACAAACCGCCCTCTCTATATCTTTTGAAGAAAGAACGATCGATCCTCTGGAGG 1320
Db 3294 CCCAGCTCGATGCAACAACATCTTGTATATTTTCAGAGAAAGCGAAGTATCTCTGGAGG 3353
QY 1321 GAAGGCGACATGTTCCCTTTTCTGATGGGTTAATAAATGCTCCAAATGATTTGACCCA 1380
Db 3354 GAAGGCGACATGTTCCCTTTTCTGATGGGTTAATAAATGCTCCAAATGATTTGACCCA 3413
QY 1381 GTTCCAAATACAAAACAACATCTTTGATGAATCCCTCTAGTCTGTGGTCCCTCGCTCAGGAA 1440
Db 3414 GTTCCAAATACAAAACAACATCTTTGATGAATCCCTCTAGTCTGTGGTCCCTCGCTCAGGAA 3473
```


QY 1441 GATCAATGCTCCCAATATTAGTTAACTTTATCTTATTTCTCTAATAATAATGAG 1500
Db 3474 GATCAATGCTCCCAATATTAGTTAACTTTATCTTATTTCTCTAATAATAATGAG 3533
QY 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 1560
Db 3534 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 3593
QY 1561 GTTCAGAGAGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 3594 GTTCAGAGAGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 3653
QY 1621 GAAGGACTTTACACTGCTGTTTTTAATAAATAACAAACAAATTTGGTCTCGAGGTTGAGG 1680
Db 3654 GAAGGACTTTACACTGCTGTTTTTAATAAATAACAAACAAATTTGGTCTCGAGGTTGAGG 3713
QY 1681 CGTCTAGCCAACTCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCAACAATGAGAA 1740
Db 3714 CGTCTAGCCAACTCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCAACAATGAGAA 3773
QY 1741 AGAATCTTCTTAATAATAGACATGCTTATGACTTCTTCTCAACAAGATGGGAGGA 1800
Db 3774 AGAATCTTCTTAATAATAGACATGCTTATGACTTCTTCTCAACAAGATGGGAGGA 3833
QY 1801 ACATGCAAGTCTTGGACCTGATTTGATCGGATAGAGACTTTGTCCAAAATATT 1860
Db 3834 ACATGCAAGTCTTGGACCTGATTTGATCGGATAGAGACTTTGTCCAAAATATT 3893
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAGAGGGGACTGGTTGGGCTG 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAGAGGGGACTGGTTGGGCTG 3953
QY 1921 GGTGTAATGAGTGGACATCGACTGGGGTGTCTTACTAATTTGGGCAATTTTGTCTACTA 1980
Db 3954 GGTGTAATGAGTGGACATCGACTGGGGTGTCTTACTAATTTGGGCAATTTTGTCTACTA 4013
QY 1981 TTATCCATAGTGTCTTGAATGCTCTATCCTGATTTTGTGATCTTAAATATATC 2040
Db 4014 TTATCCATAGTGTCTTGAATGCTCTATCCTGATTTTGTGATCTTAAATATATC 4073
QY 2041 GGATAA 2046
Db 4074 GGATAA 4079

RESULT 4

US-10-267-322-1
; Sequence 1, Application US/10267322
; Publication No. US20030152590A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Puenko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143D
; CURRENT APPLICATION NUMBER: US/10/267,322
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/336,910, and US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-10-267-322-1

Query Match 74.0%; Score 1514.8; DB 16; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;

Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTTCTGATTTAGTCTCATTTTATTCAAAGTATAAAATCTCTCCCT 60
Db 5940 ATGAAGACCATATTTTCTGATTTAGTCTCATTTTATTCAAAGTATAAAATCTCTCCCT 5999
QY 61 GTTTTAGAAATTTGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTCTCCGGAAAC 120
Db 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAACT 6059
QY 121 CTCCAAAGACAGAAAGATTTTCACTGATGGGATTTTACATGAGTGGCAAAAGATTTGCT 180
Db 6060 CTCCAAAGACAGAAAGATTTTCACTGATGGGATTTTCACTGAGTGGCAAAAGATTTGCT 6119
QY 181 GATTCCCTTTTGGAGCAATCTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 240
Db 6120 GATTCCCTTTTGGAGCAATCTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 6179
QY 241 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAGTGTAAACAGCCCT 300
Db 6180 GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAGTGTAAACAGCCCT 6239
QY 301 TCTGGAAAATCCTTTGCTGCTGATCTCCCGAGTAATATCCGCGATTACCTTAAATGTAAA 360
Db 6240 TCTGGAAAATCCTTTGCTGCTGATCTCCCGAGTAATATCCGCGATTACCTTAAATGTAAA 6299
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTTCATGCAAGGGGATTTGCCCTCCATTGTTGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTTCATGCAAGGGGATTTGCCCTCCATTGTTGG 6359
QY 421 GGGGCAATTTTCTGATGATCGGGTGGCTCTTACAAATATGATACCGAGGCAAGGTCTTC 480
Db 6360 GGGGCAATTTTCTGATGATCGGGTGGCTCTTACAAATATGATACCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAAAGACAGTTTCAAGAAATGATTTTTTCT 540
Db 6420 ACTGAAGAAATATAGCAGCTATGATTTGTTAAAGACAGTTTCAAGAAATGATTTTTTCT 6479
QY 541 AGSCAAGACAAAGGTATTCGTCAATGAACTTGACCTCCCAATAAATATTTGGACAAAGC 600
Db 6480 CGSCAAGACAAAGGTATTCGTCAATGAACTTGACCTCCCAATAAATATTTGGACAAAGT 6539
QY 601 AGCAATGAAACGACAGAAATGATACGGGATGTTTTTGGCATCTCTCCAAAGAAATACAACTCC 660
Db 6540 AGTAAACGAAACGAAACGAAATGATACGGGATGTTTTTGGCGCTCTTCAAGAAATACAACTCT 6599
QY 661 ACAACAAATCAACATGCTCCCTCATCTCTTAAACCTCCCTCCGCAAGTAACTCCG 720
Db 6600 ACAAGAAACAAACATGCTCCCTCATCTCTTAAACCTCCCTCCGCAAGTAACTCCG 6659
QY 721 AGCATTCACCTCTACAAATCTCAAAATTTAAATCTGCTAAATCTGGAATCTATGAACCCCAAGT 780
Db 6660 GAGTCAAACTCACAGCACCCCACTGATGCCCAAACTCAATACCAAGGACCCCAAGC 6719
QY 781 AGCGAATGAGGACCTTATGATTTTCGGCTCAGGATCTCGAGAACAGGGGCCCCCACA 840
Db 6720 AGTATGATGAGGACCTTCGCAACATCGGCTCAGGGTCCGAGAAACGAGAAACCCCAACA 6779
QY 841 ACTCTTAATGATGATCACTGAACAGAAACATCTGTCACAAATATTTGTCCTCTCTCACTA 900
Db 6780 ACTTCTGATCGGTTCACCAAGCAAGGGCTTTTCATCAACAAATGCCCACCTCCCTCAACA 6839
QY 901 CATCCAAAGACCTTCACAAACATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAAC 960
Db 6840 CAACCAAGCAGCCACAGCAAGGAGGAAACACAAACCAATTTCCCAAGATGCTGTGACT 6899
QY 961 GAGCACAATGGAACCGGACCCCAACAAACAGCAACCGTCTCTCAACAAATCTCAATATACA 1020
Db 6900 GAACTAGCAAAAATAACACAACTGCAACACCGTCCATGCCCTCTATAACACTTACCACA 6959
QY 1021 ACTCCCAACCTTAAACACTCTCAAGTACAACTCTGATGACTCTCTCCCTCCCAACCGGCAAC 1080
Db 6960 ATCTCTACTAACCAACCTTCCTCCAAACACAACTTCTGCACTCTCTCTGCAACCAATTAACAAAC 7019

```
QY 1081 ATCAACAAATATGATATACACAACTGTAAGTACAGAGAAACGGAACAAACCAATGCTCAGTTG 1140
Db |||||
QY 7020 ACCACCAATGACAAACACACAGAGCAACATCACTGAAATGAGCAAAACCAAGTCCGCCCTCG 7079
Db |||||
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACACAGAGCAAGACACCAACAGACACACC 1200
Db |||||
QY 7080 ATAACAACCTGCTCCCAACGGAATCCCAACAGCAAAAGAGCACCAGACGACAAA 7139
Db |||||
QY 1201 AACATCATCATGACGACATCAGATATAACAAGCAAAACCCACCAAAATTTCTTCGGAT 1260
Db |||||
QY 7140 GGCCCCGACACACGGGACCAACACAGCAAAATGAGCATTTTCACAGTCCTCCCCCACC 7199
Db |||||
QY 1261 TCTAGTCCGCAACCCGCCCTCTATATATCTTTAGAAAGAAACCGATCGATCCTCTGGAGG 1320
Db |||||
QY 7200 CCCAGCTCGATGCAACACATCTTGTATATTTTCAAGAAAGCGAAGTATCCTCTGGAGG 7259
Db |||||
QY 1321 GAAGCGCATGTTCCCTTTCTGATGGGTATATAATGCTCCAAATGATTTTGACCCA 1380
Db |||||
QY 7260 GAAGCGCATGTTCCCTTTCTGATGGGTATATAATGCTCCAAATGATTTTGACCCA 7319
Db |||||
QY 1381 GTTCCAAATACAAAACCAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 1440
Db |||||
QY 7320 GTTCCAAATACAAAACCAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 7379
Db |||||
QY 1441 GATCAACATGCTCCGCCCAATATAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Db |||||
QY 7380 GATCAACATGCTCCGCCCAATATAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 7439
Db |||||
QY 1501 AACCTGCTACTCTGGAGAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db |||||
QY 7440 AACCTGCTACTCTGGAGAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 7499
Db |||||
QY 1561 GTTCAGGAGGATGACCTGCGCGAGGCTCAGTTGCGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db |||||
QY 7500 GTTCAGGAGGATGACCTGCGCGAGGCTCAGTTGCGATACCGTTTTTTTGGCCCTGGAAAT 7559
Db |||||
QY 1621 GAAGGACTTTACACTGCTGTTTTTAATAAATAACAAACAAATTTGGTCTGCAAGTTGAGG 1680
Db |||||
QY 7560 GAAGGACTTTACACTGCTGTTTTTAATAAATAACAAACAAATTTGGTCTGCAAGTTGAGG 7619
Db |||||
QY 1681 CGCTAGGCAATCAAACTGCGCAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Db |||||
QY 7620 CGCTAGGCAATCAAACTGCGCAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 7679
Db |||||
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAGATGGGGGGA 1800
Db |||||
QY 7680 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAGATGGGGGGA 7739
Db |||||
QY 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAAGACTTGTCCAAAATAATT 1860
Db |||||
QY 7740 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAAGACTTGTCCAAAATAATT 7799
Db |||||
QY 1861 TCAGAGCAAAATGACAAATTAATAAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db |||||
QY 7800 TCAGAGCAAAATGACAAATTAATAAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 7859
Db |||||
QY 1921 GGTGCTTAATGGTGACATCCGACTGGGGTGTCTTACTTAACCTGGGCAATTTGCTACTA 1980
Db |||||
QY 7860 GGTGCTTAATGGTGACATCCGACTGGGGTGTCTTACTTAACCTGGGCAATTTGCTACTA 7919
Db |||||
QY 1981 TTATCCATAGCTGCTCTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 2040
Db |||||
QY 7920 TTATCCATAGCTGCTCTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 7979
Db |||||
QY 2041 GGATAA 2046
Db |||||
QY 7980 GGATAA 7985
Db |||||
```

RESULT 5
US-10-066-506A-9
; Sequence 9, Application US/10066506A

```
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 9
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Marburg virus strain Musoke
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-9
```

Query Match 72.6%; Score 1485.6; DB 15; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;

```
QY 1 ATGAAGACCAATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db |||||
QY 10 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATCAAGGACNAAAANTCTCCC 69
Db |||||
QY 61 GTTTTAGAAATTTGTAGTAAACAGCCAACTCAAGATGTAGATTAGTGTGTCGGAAC 120
Db |||||
QY 70 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTGGTATGTCGGAAC 129
Db |||||
QY 121 CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGCAAAAGTTGCT 180
Db |||||
QY 130 CTCAGAAGACAGAAAGCGTCATCTGATGGGATTTACACATGAGTGGGCAAAAGTTGCT 189
Db |||||
QY 181 GATTCCCTTTTGAAGCATCTTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db |||||
QY 190 GATTCCCTTTTGAAGCATCCNAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 249
Db |||||
QY 241 GTTCAGTATACGGAAGGAGAAAGCAAAACATGTTTCAATATAAGTGTAAACAGCCCT 300
Db |||||
QY 250 GTTCAGTACACAGAGGGGAGAAAGCAAAACATGCTACATATAAGTGTAAACAGATCCC 309
Db |||||
QY 301 TCTGGAATAATCTTGTCTGGATCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 360
Db |||||
QY 310 TCTGGAATAATCTTGTCTGGATCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 369
Db |||||
QY 361 ACTGTTCAATATTCAGGTCAAAACCTCATGCACAGGGGATTCGCCCTCCATTTGTGG 420
Db |||||
QY 370 ACTATCATCATATTCAGGTCAAAACCTCATGCACAGGGGATTCGCCCTTCATTTATGG 429
Db |||||
QY 421 GGGCAATTTTCTGATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGGCTTC 480
Db |||||
QY 430 GGAGCAATTTTCTGATGTCGCAAT-----ATGTACCGAGGCAAGGCTTC 477
Db |||||
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATGAAGACAGTTTCACAGATGATTTTTCT 540
Db |||||
QY 478 ACTGAAGGGAACATAGCAGCTATGATTTGTTCAATGAAGACAGTGCACAAAATGATTTTCG 537
Db |||||
QY 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTGGACAAGC 600
Db |||||
QY 538 CGGCAAGGACAGGTTATCGTCATATGATCTGATCTTACTTAATATAATTTGACAAGT 597
Db |||||
QY 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTTTGCGATCTCTCCAAAGATACAACTCC 660
Db |||||
QY 598 AGTAACGGAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 657
Db |||||
QY 661 ACAACATCAAACTGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
Db |||||
QY 658 ACNAAAGAACCAAAATGCTGCTCTCGTCCAAAATACCTCCACCTGCCCCCAGCCCGTCCG 717
Db |||||
```

Qy	721	AGCATTCACTCTTACAAATATCTCAAAATTAAATCTGTCTAAATCTTGAACTATGAAACCCAAAGT	780
Db	718	GAGATCAAACTCTCAACGACACCCCAACTGATGTGCACCAAACTCTAATACCGAGCCCAAGC	777
Qy	781	AGCGAGATCAGGACCTTATGATTTCCGGCTCAGGATCTCGAGAACAGGGCCCCCACA	840
Db	778	AGTGATGATGAGGACCTTCGCACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA	837
Qy	841	ACTCTTAAATAGTACTGTACGACGAAACAATCGTCAACAAATATTGTGCCACTCTCTCACTA	900
Db	838	ACTTCTGATGGGTACCAAGCAAGGGCTTTTCATCACAATGCAACCCACTCTCCTCAACA	897
Qy	901	CATCCAAAGCACTTCAACAATGAGCAAAAACAGTACGAATCTTCCCGACATGCTGTAACT	960
Db	898	CAACCAAGCACGCCACAGACGAGGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT	957
Qy	961	GAGCACAAATGGAAACCGACCCAAACAACAACACGACGACGCTCTCTCAACATATCTTAATACA	1020
Db	958	GAATGAGCAAAATAACAACATGCAACAACGGTCCATGCCCCCTCATAAACATCACACA	1017
Qy	1021	ACTCCCACTTATAACACTCTCAAGTACAACTCAGTACTCTTCTCCCTCCAAACCCGCAAC	1080
Db	1018	ATCTCTACTAACACACTCTCCAAACAACTTCAGCACTCTCTCTGCACCACTTACAAAAC	1077
Qy	1081	ATCACCAATATGATACAAACGTGAATCTAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
Db	1078	ACCACCAATGACAAACAACAGAGCACAACTCACTGNAATGAGCAAAACAGTGCCTCCCTCG	1137
Qy	1141	AACACAACCTTAGATCCAAACAGAAAATCCCAACACAGGACAAGACACCAACAGACAAC	1200
Db	1138	ATAACAACCTTGCTCCAAAGGAAATCCCAACACAGCAAAAGAGCACCGAGCGAANAATA	1197
Qy	1201	AACATCATATGACGACATCAGATATATAACAGCAAAACACCCCAAAATCTTCTCGGAT	1260
Db	1198	GGCCCCCAACAACGGCACCAACAACAGCAAAATGAGCATTTTCAACCACTCTCCCCCAAC	1257
Qy	1261	TCTAGTCCGACACCCCGCTCTCTATATATCTTTAGAAGAAACGATCGATCTCTCGAGG	1320
Db	1258	CCCAGCTCGAATGCACAACATCTTGATATATTTCAAGAAGAGCGATCGATCTCTCGAGG	1317
Qy	1321	GAAGGACATGTTTCCCTTTTCTGGATGGGTAAATAAATGCTCCAAATTGATTTTGACCCA	1380
Db	1318	GAAGCGACATGTTTCCCTTTTCTGGATGGGTAAATAAATGCTCCAAATTGATTTTGACCCA	1377
Qy	1381	GTTCCAAATACAAAAACAATTTTGATGAATCTCTAGTTCTGGTGCTCGGCTGAGGAA	1440
Db	1378	GTTCCAAATACAAAAACAATTTTGATGAATCTCTAGTTCTGGTGCTCGGCTGAGGAA	1437
Qy	1441	GATCAACATGGCTCCCCCAATATAGTTTAACTTTATCTTATTTTCTTAATAATAAATGAG	1500
Db	1438	GATCAACATGGCTCCCCCAATATAGTTTAACTTTATCTTATTTTCTTAATAATAAATGAG	1497
Qy	1501	AACACTGCCTACTCTGGAGAAAATGAGAAATGATGTGATGCAGAGTTTAAAGAAATTTGAGC	1560
Db	1498	AACACTGCCTACTCTGGAGAAAATGAGAAATGATGTGATGCAGAGTTTAAAGAAATTTGAGC	1557
Qy	1561	GTTCAGGAGGATGACTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGGAATT	1620
Db	1558	GTTCAGGAGGATGACTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGGAATT	1617
Qy	1621	GAAGGACTTTACATGCTGTTTTTAAATAAATCAAAACAATTTGGTCTCCAGGTTGAGG	1680
Db	1618	GAAGGACTTTACATGCTGTTTTTAAATAAATCAAAACAATTTGGTCTCGAGGTTGAGG	1677
Qy	1681	CGTCTAGCCAACTCAAACTGCCAAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1740
Db	1678	CGTCTAGCCAACTCAAACTGCCAAATCTCTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1737
Qy	1741	AGAACATTCTCTTAAATCAATPAGACATGCTATTCATCTTTCTACACAAGATGGGAGGA	1800
Db	1738	AGAACATTCTCTTAAATCAATPAGACATGCTATTCATCTTTCTACTCACAAGATGGGAGGA	1797

Qy	1801	ACATGCAAAAGTCTTGGA	CTGATTGTTGCATCGGATAGAA	GACTTGTCCAAAAATATT	1866
Db	1798	ACATGCAAAAGTCTTGGA	CTGATTGTTGCATCGGATAGAA	GACTTGTCCAAAAATATT	1857
Qy	1861	TCAGAGCAAAATTCAC	CAAAATTTAAAAAGGACGAA	CAAAAAAGAGGGGACTG	1920
Db	1858	TCAGAGCAAAATTCAC	CAAAATTTAAAAAGGACGAA	CAAAAAAGAGGGGACTG	1917
Qy	1921	GGTGGAATAATGGTG	GACATCCGACTGGGGTGTTCTTA	CTAACTTGGGCAATTTTGCTACTA	1980
Db	1918	GGTGGAATAATGGTG	GACATCCGACTGGGGTGTTCTTA	CTAACTTGGGCAATTTTGCTACTA	1977
Qy	1981	TTATCCATAGCTGTCT	TGATTGCTCTATCCTCTGATTTTG	CTGATCTTTACTAAATATATC	2040
Db	1978	TTATCCATAGCTGTCT	TGATTGCTCTATCCTCTGATTTTG	CTGATCTTTACTAAATATATC	2037
Qy	2041	GGATAA	2046		
Db	2038	GGATAA	2043		
RESULT 6					
US-10-811-353-2					
; Sequence 2, Application US/10811353					
; Publication No. US20050112098A1					
; GENERAL INFORMATION:					
; APPLICANT: McGray, Paul B.					
; APPLICANT: Sanders, David A.					
; APPLICANT: Jeffers, Scott A.					
; APPLICANT: Davidson, Beverly L.					
; APPLICANT: Sinn, Patrick L.					
; TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE					
; FILE REFERENCE: 290.00670120					
; CURRENT APPLICATION NUMBER: US/10/811,353					
; CURRENT FILING DATE: 2004-03-26					
; NUMBER OF SEQ ID NOS: 16					
; SOFTWARE: PatentIn version 3.2					
; SEQ ID NO 2					
; LENGTH: 2046					
; TYPE: DNA					
; ORGANISM: Marburg virus glycoprotein					
US-10-811-353-2					
Query Match 72.0%; Score 1473.2; DB 21; Length 2046;					
Best Local Similarity 82.5%; Pred. No. 0;					
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;					
Qy	1	ATGAAGACCATATATTTT	CTGATTAGTCTCATTTTAA	TCCAAAGTATAAAACTCTCCCT	60
Db	1	ATGAAGACCATATATTTT	CTGATTAGTCTCTTTTAT	CAGTCTTATCTTAATCCAAGGGATAAAACTCTCCCT	60
Qy	61	GTTTTAGAAATTCGTAG	TACAGCCACCTCAGAGTGTAGAT	TTCAGTGTCTCCGGAACT	120
Db	61	ATTTTAGAGATGAGCTG	ATAGTAACTCAACCCCAAAATGTG	ATTCGGTATGTCCTCGGAAT	120
Qy	121	CTCCAAAGACAGAAATG	TTTCATCTCATGGGATTTTAC	CTGAGTGGGCAAAAAGTTGCT	180
Db	121	CTCCAGAGACAGAAATG	TTTCATCTCATGGGATTTTAC	CTGAGTGGGCAAAAAGTTGCT	180
Qy	181	GATTCCCTTTGGGAAG	CAATTTAAACGATGGGCTTTT	CAGGACAGGTGTCTCCCAAGAC	240
Db	181	GATTCCCTTTGGGAAG	CAATTTAAACGATGGGCTTTT	CAGGACAGGTGTCTCCCAAGAC	240
Qy	241	GTTGAGTATACGAGAG	GAAGCAACCAATGTTTACA	TATTAAGTGTAAACAGCCCT	300
Db	241	GTTGAGTATACGAGAG	GAAGCAACCAATGTTTACA	TATTAAGTGTAAACAGCCCT	300
Qy	301	TCTGGAAAACTCTTG	CTGCTGGATCTCTCCCA	GTAATATCCCGGATTAACCTTAATGTAAA	360
Db	301	TCTGGAAAACTCTTG	CTGCTGGATCTCTCTAC	CAATATCCCGGATTAACCTTAATGTAAA	360
Qy	361	ACTGTTTCATATATT	TCAAGGTCAAAACCCCT	CATGCAAGGGGATTCGCCCTCCCAATTTGTGG	420

Dd 361 ACTATCCATCATATATTCAAAGGTCAAAACCCCTCATGTCGCGAAGGGATGCGCCCTCCATTTGTGG 420
Qy 421 GGGGCAATTTTCTGTATGATCGCGTTCCTCTCAACAACATGTATACCGAGGCAAGTCTTC 480
Dd 421 GAGCAATTTTCTGTATGATCGCATTCGCTCCACAACAATGTACCGAGGCGAGTCTTC 480
Qy 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTCACAGAATGATTTTCT 540
Dd 481 ACTGAAGGGAACATAGCAGCTATGATTTGTTAATAAGACAGTTCACAAAATGATTTCTCG 540
Qy 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATATATATGACAGC 600
Dd 541 AGGCAAGGACAGGTTATCGTCACATGAATCGCTCAACATGATCTGACTTCTAATAAATATATGACAGT 600
Qy 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAATCC 660
Dd 601 AACAAATGGAACACAACCGAATGACACTGGATGCTTCGGTCTCTCAAGAAATACAATCC 660
Qy 661 ACAAAATCAAAATGCGCTCCATCTCTTAAACCTCATCCCTGCGCCACAGTACTCCG 720
Dd 661 ACGAAGAAATCAAAATGCTCCGTCCAAATATACCTCACCCTGCCCCACAGCCCGTCCA 720
Qy 721 AGCAATTCCTACAAATACTCAAAATTAATCTGCTAAATCTGCTAAATCTGAACTATGAACCAAGT 780
Dd 721 GAGATCAAAACCCACAAGCACCCCAACTGATGCCACCACTCAACACCAAGACAGCCCAAC 780
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Dd 781 AATGATGATGAGGACCTCATAACTCCGTTTCAGGGTCGGAGAACAGAACCCCTATACA 840
Qy 841 ACTCTTAATGATGACCTGAAACAGAAACAATCGTCAACAATATGTTCCACTCTCTTCACTA 900
Dd 841 ACTTCAGATGGGTCACCTAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACC 900
Qy 901 CATCAAGCACTCAACATGAGCAGCAACAGTACGATCTCCGACATGCTGTAAC 960
Dd 901 CAACCAAGCAGCCACAGCAAGAGGAAACAACAGACACCACTTCCCAAGGTAAGTGTGACT 960
Qy 961 GAGCAAAATGGAACCGACCCCAACACACAACAGCAACGCTCCTCAACAATACTAATAACA 1020
Dd 961 GAACCCACAAACCAACAACAGGCAACACCGTCCATGCGCCGCCCAACACCACTGCA 1020
Qy 1021 ACTCCCACTATAACACTCTCAAGTACAACTCAGTACTCTCTTCCCTCCCAACCGCAAC 1080
Dd 1021 ATCTCTACTAAACACCTCCAAAGAACTTTCAGCACCTCTCTGTATCACTACAAAC 1080
Qy 1081 ATCAACCAATATGATACACACGTTGAATAGCAGAAAGGCAACCAACCAATGCTCAGTTG 1140
Dd 1081 ACCCAATATGACACACAGGACAGCCACTGAAATGAAACCAACAGTGCCTCCCTCG 1140
Qy 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGGACAGACCAACAGCAGCAAC 1200
Dd 1141 AAACAACCTCTCCCAACAGGAATCTTACCAGCAAGAGCACTTAAACAACGAA 1200
Qy 1201 AACATCATGACGACATCATAGATATAACAGCAAAACACCCCAAAATCTTCTCCGAT 1260
Dd 1201 GCGCCCAACCAACCGCACCAAAATATGACAAATGGCATTTTAAACAGTCCCTCCGCCACC 1260
Qy 1261 TCTAGTCGCAACACCGCCCTCTATATACCTTGAAGAGAACGATCGATCTCTGGAG 1320
Dd 1261 CCCAACCCGACACACAACATCTTGTATATTTTCAAGAAAGAACGAGTATCTCTCGAGG 1320
Qy 1321 GAAGGCGACATGTTCCCTTTCTGATGGGTTAATAAATGCTCCCAATGATTTTGACCCA 1380
Dd 1321 GAAGGCGACATGTTCTTTCTGACGGGTTAATAATGCTCCCAATGATTTTGATCCA 1380
Qy 1381 GTTCAAAATACAAAACCAATCTTTGATGAATCTCTAGTCTGCGCTCCGCTGAGAA 1440
Dd 1381 GTTCAAAATACAAAGACGATCTTTGATGAATCTTCTAGTCTGCGCTGAGAA 1440
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTCTTAATAATAAATGAG 1500
Dd 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTCTTAATAATAAATGAA 1500

Qy 1501 AACACTCCCTACTCTCGAGAAAAATGAGAATGATTTGTATGTCAGAGTTAAAGAAATTTGGAGC 1560
Dd 1501 AACACTCCCTACTCTCGAGAAAAATGAGAACGATTTGTATGTCAGAGTTAAAGAAATTTGGAGC 1560
Qy 1561 GTTCAGAGGATGACCTGGCCGAGGCTCAGTTGGATACCGTTTGGCCCTCGGAAT 1620
Dd 1561 GTTCAGAGGATGACCTGGCCGAGGCTCAGTTGGATACCGTTTGGCCCTCGGAAT 1620
Qy 1621 GAAGCACTTTACACTGCTGTTTAAATTAATAAATCAAAACAATTTGGTCTCGCAGTTGAGG 1680
Dd 1621 GAAGCACTTTATCTGCTGTTTAAATTAATAAATCAAAACAATTTGGTCTCGCAGTTGAGG 1680
Qy 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGAA 1740
Dd 1681 CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGAA 1740
Qy 1741 AGACATTCCTTAAATCAATAGACATGCTATTGCACTTTCTACTCACAAGATGGGAGGA 1800
Dd 1741 AGGCAATTTCTTAAATTAATAGACATGCCATTCGACTTTCTACTCACAAGATGGGAGGA 1800
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATTT 1860
Dd 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAGGAATATTT 1860
Qy 1861 TCAGAGCAAAATTTGACCAAAATTAATAAAGGACGAAACAAAAGAGGGGACTGGTTGGGCTCTG 1920
Dd 1861 TCGGAAACAATTTGACCAAAATTAATAAAGAGTGAACAAAAGAGGGGACTGGTTGGGCTCTA 1920
Qy 1921 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTACTACTTGGGCAATTTTGGCTACTA 1980
Dd 1921 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTACTACTTGGGCAATTTTGGCTACTA 1980
Qy 1981 TTATCCATAGCTGCTTGTGATTTGCTATCCTGATTTGTCTGATCTTTTACTAAATATATC 2040
Dd 1981 TTATCCATAGCTGCTTGTGATTTGCTATCCTGATTTGTCTGATTTTACTAAATATATC 2040
Qy 2041 GGATAA 2046
Dd 2041 GGGTAA 2046

RESULT 7
US-10-353-856-27
; Sequence 27, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawasaka, Yoshihiro
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Bae
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Marburg virus
US-10-353-856-27

Query Match 72.0%; Score 1473.2; DB 17; Length 19112;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Dd 5940 ATGAAGACCATATGCTCTTTATCAGTCTTATCTTAATCCAGGATATAAACTCTCCCT 5999


```
Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAATAAAGGACGAAACAAAAGAGGGGACTGGTTGGGGTCTG 3953
Qy 1921 GGTGGTAAATGGTGACATCCGACTGGGGT 1950
Db 3954 GGTGGTAAATGGTGACATCCGACTGGGGT 3983

RESULT 9
US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE OF INVENTION: Filovirus Infection in Primates
; CURRENT FILING DATE: 2004-03-26
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US/10491121
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 8256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Marburg GP (dTM)
US-10-491-121-32

Query Match 69.3%; Score 1418.8; DB 20; Length 8256;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1618; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTCTGATAGTCTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Db 1431 ATGAAGACCATATTTCTGATAGTCTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT 1490
Qy 61 GTTTTAGAATTTGTAGTAAAGCCAACTCAAGATGTAGATTCAAGTGTCTCCGGAACT 120
Db 1491 ATTTTAGAGTAGCTAGTAAATTAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAACT 1550
Qy 121 CTCCAAAAGACAGAGATGTTTCATCTGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 1551 CTCAGAGACAGAGAGCTCCATCTGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 1610
Qy 181 GATTCCTCTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 1611 GATTCCTCTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 1670
Qy 241 GTTGAGTATACGAGAGGAGAGGACCAAAACATGTTACATATATAAGTAAACAGACCT 300
Db 1671 GTTGAGTATACGAGAGGAGAGGAGCAAAACATGTTACATATATAAGTAAACAGACCT 1730
Qy 301 TCTGGAATATCTTCTGCTGATCCTCCAGTAAATATCCGCGATTTACCTAAATGTAAA 360
Db 1731 TCTGGAATATCTTCTGCTGATCCTCCAGTAAATATCCGCGATTTACCTAAATGTAAA 1790
Qy 361 ACTGTTTCAATCATTAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCATTTTGG 420
Db 1791 ACTATCCATCATTAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCATTTTGG 1850
Qy 421 GGGCATTTTCTGATGATCGCTTGCCTCTACAAATGATACCGAGGCAAGGTCTTC 480
Db 1851 GGAGCATTTTCTGATGATCGCTTGCCTCTACAAATGATACCGAGGCAAGGTCTTC 1910
```

```
Qy 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGATGATTTTTTCT 540
Db 1911 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGATGATTTTTTCTCG 1970
Qy 541 AGGCAAGGACAAAGTTATCTGCATGAACTTGACCTCCCAATATAATATTGGACAAGC 600
Db 1971 CGGCAAGGACAAAGGTTACCGTCAATAGAACTGACTTCTACTATAATAATATTGGACAAGT 2030
Qy 601 AGCAATGAAACGCAGAGAAATGATACGGGATGTTTGGGATCTCTCCCAAGATATACTCC 660
Db 2031 AGTAACGGAAACGAAACGAATGACACTGGATGTTTGGGCTCTCTTCAAGAATACTAATTTCT 2090
Qy 661 ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
Db 2091 ACAAAGAAACCAACATGTGCTCCGTCCTCAAAATACCTCCACCACTGCCACAGACCCGCTCG 2150
Qy 721 AGCATTCCTCTACAAATCTCAAAATTAATACTGCTAAATCTGGAATCTATGAAACCCCAAGT 780
Db 2151 GAGATCAAACTCAACAGACCCCACTGATGCGCAACCAACTCAATACCGAACCCCAAGC 2210
Qy 781 AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 2211 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 2270
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTGTCCACTCTCTTCACTA 900
Db 2271 ACTTCTGATGCGGTCAACAAAGCAAGGGCTTTTCATCAACAATGCCCACCACTCCCTCAACA 2330
Qy 901 CATTCCAGACCTCTACAACTGAGCAAAACAGTAGCAATCTCTCCGACATGCTGTAACCT 960
Db 2331 CAACCAAGCACGCCACAGAGGAGGAAACAACAACCAACCAATCTCCCAAGATGCTGTGACT 2390
Qy 961 GAGCACAATGGAAACCGACCCCAACAACAACAGCAACCGTCTCTCAACAATACTAATAACA 1020
Db 2391 GAATAGCAAAATAACAACAATGTCACACCGTCCATGCCCTCATACACTTACCACA 2450
Qy 1021 ACTCCCACTTATTAACACTCTCAAGTACAACCTCAGTACTCTCTCCCTTCCAAACCCGCAAC 1080
Db 2451 ATCTCTACTAAACAACACTCTCAACAACAACCTCTCAGCACTCTCTCTGCAACCAATTAACA 2510
Qy 1081 ATCAACCAATTAATGATACAAACGTTAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Db 2511 ACCACCAATGACACACACAGAGCAACATCACTGAAAATGAGCAACCGAGTGGCCCTCTG 2570
Qy 1141 AACACAACCTTAGATCCAAACAGAAAATCCCAACACAGAGCAAGACACCAACAGCAACAAC 1200
Db 2571 ATAACAACCTGCTCTCAACGGGAATCCCAACACAGAGCAAGACACCAAGAGCAACAGAGCA 2630
Qy 1201 AACATCATATGACGACATCAGATATTAACAGCAAAACACCCCAACAAATTTCTTCCGGAT 1260
Db 2631 GGCCCCGCCCAACAGGCAACCAACAACGACAAATGAGCAATTTACAGCTCTCTCCGCCAC 2690
Qy 1261 TCTAGTCCGACAAACCCGCCCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAG 1320
Db 2691 CCCAGCTCGATGCAACAACATCTTGTATATTTTCAGAAAGAAACGAGATATCTCTGGAG 2750
Qy 1321 GAAGCGACATGTTCCCTTTTCTGATGGGTTAATAATGCTCCAAATGATTTGTTGACCCA 1380
Db 2751 GAAGCGACATGTTCCCTTTTCTGATGGGTTAATAATGCTCCAAATGATTTGTTGACCCA 2810
Qy 1381 GTTCCAAATACAAAACAATTTTGTATGAATCTCTTCTAGTCTGCTGCTCGGCTGAGGAA 1440
Db 2811 GTTCCAAATACAAAACAATTTTGTATGAATCTCTTCTAGTCTGCTGCTCGGCTGAGGAA 2870
Qy 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTATATAATAAGTAG 1500
Db 2871 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTATATAATAAGTAG 2930
Qy 1501 AACACTGCTACTCTCGAGGAAATGAGAAATGATTTGTATGATGAGTAAAGAAATTTGGAGC 1560
Db 2931 AACACTGCTACTCTCGAGGAAATGAGAAATGATTTGTATGATGAGTAAAGAAATTTGGAGC 2990
```



```
QY 1561 GTTCAGGAGGATGACCTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
Db 2991 GTTCAGGAGGATGACCTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 3050
QY 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTGCAGGTGAGG 1680
Db 3051 GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTGCAGGTGAGG 3110
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAACTCTTTATTGAGAGTCAACAACCTGAGGAA 1740
Db 3111 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAACTCTTTATTGAGAGTCAACAACCTGAGGAA 3170
QY 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTTACTCACAAGATGGGAGGA 1800
Db 3171 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTTACTCACAAGATGGGAGGA 3230
QY 1801 ACATGCAAGTCTTGGACCTGATTGTTGCATCCGGATAGAGACTTGTCCAAAATATT 1860
Db 3231 ACATGCAAGTCTTGGACCTGATTGTTGCATCCGGATAGAGACTTGTCCAAAATATT 3290
QY 1861 TCAGAGCAAAATTGACCAAAATTAATAAGGACGAACAAAAGAGGGGACTGTTGGGGTCTG 1920
Db 3291 TCAGAGCAAAATTGACCAAAATTAATAAGGACGAACAAAAGAGGGGACTGTTGGGGTCTG 3350
QY 1921 GGTGTAATGGTGACATCCGACTGGGT 1950
Db 3351 GGTGTAATGGTGACATCCGACTGGGT 3380

RESULT 10
US-10-066-506A-5
; Sequence 5, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 5
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 1
; OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 64.5%; Score 1319.6; DB 15; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

QY 1 ATGAGACCATATATTTCTGATTAGTCTCAATTTTAATCCAAAGTATATAAACTCTCCCT 60
Db 1 ATGAGACCATATGTTTCTTATAGTCTTATCTATTATCTTTAATTCAGGGACAAAATCTCCCC 60
QY 61 GTTTTAAAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
Db 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120
QY 121 CTCAAAAAGACAGAAGATGTTCACTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAAGACAGAAGAGCTCCATCTGATGGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGCATCTTAAAGATGGGTTTTCCAGACAGGTTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTGGAGCATCTTAAAGATGGGTTTTCCAGACAGGTTTCTCCCAAGAAC 240
```

```
QY 241 GTTGAGTATACGGAAGGAGAAGAGCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 300
Db 241 GTTGAGTATACAGAGGGGGAGGAAGCCAAACATGCTTACAATAAAGTGTAAACAGGATCCC 300
QY 301 TCTGGAATTCCTTGTCTGCTGATCCTCCAGTAAATATCCGCGATTACCTTAATATGTA 360
Db 301 TCTGGAATTCCTTGTCTGCTGATCCTCCAGTAAATATCCGCGATTACCTTAATATGTA 360
QY 361 ACTGTTTCATCATATTTCAAGGTCMAAACCCCTCATGCACAGGGGATGCGCCTCCATTTTGG 420
Db 361 ACTATCCATCATATTTCAAGGTCMAAACCCCTCATGCACAGGGGATGCGCCTCCATTTTGG 420
QY 421 GGGCATTTTCTTGTATGATCGGTTTGCCTCTTACAACAATGTAACGAGGCAAGGTCTTC 480
Db 421 GGAGCATTTTCTTGTATGATCGCATTTGCCTCTCCACAACAATGTAACGAGGCAAGGTCTTC 480
QY 481 ACTGAAGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGATGATTTTTTCT 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTTCAATAAGACAGTTTCACAGATGATTTTTTCT 540
QY 541 AGGCAAGGACAAAGGTTATCGTCAATGAACCTTGACCTCCACCAATAAATATTGGACAAGC 600
Db 541 CGGCAAGGACAAAGGTTACCGTCAATGAATCTGACTTCTACTAATAAATATTGGACAAGT 600
QY 601 AGCAATGAAACGCGAGAGAAATGATA CGGGATGTTTTTGGCATCTCTCCAGAAATACAACTCC 660
Db 601 AGTAAACGGAACGCAACGAAATGACATCGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660
QY 661 ACAACCAATCAAACTGCCCTCCATCTTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
Db 661 ACAAAAGCAACCAATCTGCTCGTCCGTCAAAATACCTCCACACTGCCACACAGCCGCTCG 720
QY 721 AGCATTCACCTACAAATACCTCAAAATTAATCTCTAAATCTGGAACCTATGAACCCAACT 780
Db 721 GAGATCAAACTCAACAGCAACCCCAACTGATGTCACCAACTCAATACCAACGAGCCCAAGC 780
QY 781 AGCGACATGAGGACCTTATGATTTCCGGTCTCAGGATCTGGAGAACAGGGGGCCCAACA 840
Db 781 AGTGATGATGAGGACCTCGCAACATCGGGTCTCAGGGTCCGGAGAACAGGAAACCCCAACA 840
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAACTCGTCAACAAATATTGTCCACTCTTCACTA 900
Db 841 ACTTCTGATCGGTCAACCAAGCAAGGCTTTTCATCAACAAATGCCACCACTCCCTCAACA 900
QY 901 CATCCAAGCACCTCAACAACATGAGCAAAAACAGTAGTACGAATCTCTCCGACATGCTGTA 960
Db 901 CAACCAAGCACGCCACAGCAGGAGGAGGAACACACAAACCAATTCCTCAAGATGCTGTGACT 960
QY 961 GAGCAATGGAAACCGCAACCAACCAACCAACCAAGCAAGCTCTCTCAACAATACTAATAACA 1020
Db 961 GAACTAGCAAAAATAACAACTGCAACACCGTCCATGCCCTCCCTCATPAACACTACCACA 1020
QY 1021 ACTCCGACCTTATACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 1021 ATCTCTACTAACAACTCCTCAAAACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
QY 1081 ATCAACCAATATGATACAAACGTTGAACCTAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
Db 1081 ACCACCAATGACACACACAGAGCACAACTCACTGAAATGAGCAAAACAGTGGCCCTCG 1140
QY 1141 AACCAACTCTTAGATCCAAAGAAAATCCCAACCAAGGACCAAGACACCAAGCAACACC 1200
Db 1141 ATAACAACCTCTCCCAACGCGGAAATCCCAACCAAGGACCAAGGACCAAGCAACCAAGCA 1200
QY 1201 AACCATCATGACGACATCAGATATTAACAGCAAAACACCCCAACCAATTTCTTCCGAT 1260
Db 1201 GGCCCCGCCAACAGGCAACCAACACGACAAATGAGCAATTTTCAAGTCTTCCCTCCCAAC 1260
QY 1261 TCTAGTCCGACAAACCGCCCTCTTATATCTTTTAAAGAAAGAACGATCGATCTCTGGAGG 1320
Db 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTTGGAAA 1320
```



```
QY 1321 GAAGGACATGTTCCCTTTCTCGATGGGTTAATAATGCTCAATGATTGTTGACCCA 1380
Db 1321 GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAATAATGCTCAATGATTGTTGATCCA 1380
QY 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCCTGGTCCGCTCGGCTGAGGAA 1440
Db 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCGACTTAACTACTTCAACTATGAG 1440
QY 1441 GATCAACATGCTCCCAATATATAGTTTAACTTTATCTTTATTTCTTAAATATATAATGAG 1500
Db 1441 GAACAACACACTCCCGCAATATCAGTTTAACTTTCTTATTTCTCTGATAAAAAATGGA 1500
QY 1501 AACACTGCTACTCTGGAGAAAATCAGAAATGATGATGATGATGATGATGATGATGATG 1560
Db 1501 GATACTGCTACTCTGGGGAACACAGAAATGATGATGATGATGATGATGATGATGATGAGT 1560
QY 1561 GTTCAGAGAGGATGACTCGCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTGAGAGAGGAGATTTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
QY 1621 GAAGGACTTTACATGCTGTTTTTAAATTAATAATCAAAACAATTTGGTCTGAGGTTGAGG 1680
Db 1621 GAAGGACTTACTACTGCGGTTTTAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
QY 1681 CGTCTAGCAATCAAACTGCCAAATCTTTGAACTCTTATTTGAGAGTCAACAATGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTTAAATCTTTGGAGCTCTTTGTTAAGGGTCAACAACGAGGAA 1740
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTTGACTTTCTACTCAACAAGATGGGAGGA 1800
Db 1741 AGGACATTTCTTAATCAATAGGATGCAATTTGACTTTTCTTTTACAGAGTGGGGGGA 1800
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAATAATATC 1860
Db 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAATAATATC 1860
QY 1861 TCAGAGCAATTTGACCAATTAATAAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
QY 1921 GGTGTAATGCTGGACATCGACTGGGGTCTTCTACTACTTGGGATTTTCTGCTACTA 1980
Db 1921 GGTGGCAATGGTGACATCTGACTGGGGTGTCTCAACAATTTGGGATCTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTGAATGCTCTATCTCTGATTTGTCGATCTTTTACTAATAATATC 2040
Db 1981 TTAATATAGCTGCTTGAATGCTCTGCTCTGATCTGCTGATCTGCTGATCTTCACTAATAATATC 2040
QY 2041 GGATAA 2046
Db 2041 GGATGA 2046
```

RESULT 11

```
US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 42
; LENGTH: 6902
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized)
US-10-491-121-42

Query Match      40.8%; Score 834; DB 20; Length 6902;
Best Local Similarity 64.3%; Pred. No. 1.3e-231;
Matches 1251; Conservative 0; Mismatches 695; Indels 0; Gaps 0;

QY 1 ATGAGACCATATATTTTCTGATAGTCTCATTTTATCTCAAGTATATAAACTCTCCCT 60
Db 1923 ATGAAGACCACTGCTGCTTTTCATCAGGCTGATCTCTGATCCAGGGCATCAAGACCTGCCC 1982
QY 61 GTTTTGAATAATGCTAGTAACAGCAACCTCAAGATGTAGATTCAAGTGTGTGTCGGGAACC 120
Db 1983 ATCTTGAGATCGCAGCAACACACGCCCCAGAACGTTGGACAGCGTGTGAGCGGCACC 2042
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAAGTTGCT 180
Db 2043 CTGCAGAAAGACCGAGGACGTGCACCTGATGGGCTTTCACCTGAGCGGCGCAGAGGTGCC 2102
QY 181 GATTCCTCTTTGGAAGCATCTAAAGATGGGCTTTTCAGACAGAGTGTTCCTCCCAAGAAC 240
Db 2103 GACAGCCTCTGGAAGGCGCAGCAAGAGGTGGGCTTTCAAGACCGGCTGCCCCCAAGAAC 2162
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACAATGTTTACATATATAGTTTAACAGACCCCT 300
Db 2163 GTGAGGTACACCGAGGGGAGAGGCGCAAGACCTGCTTACAACATCAGCGTGCACCGACCCC 2222
QY 301 TCTGAAAAATCTTTGCTGCTGGATCTCTCCAGTAAATATCCGCGATTAATACCTTAAATGTA 360
Db 2223 AGCGCAAGAGCGCTGCTGCTGGACCTTCCCAACAACATCAGGGACTACCTTAAGTGCAG 2282
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTTCATGCA CAGGGATTTGCCCTTCATTTGFG 420
Db 2283 ACCATCCACACATCCAGGGCCAGAACCTTCA CCGCCAGGGCATCGCCCTGACCTGTGG 2342
QY 421 GGGCATTTTCTGTATGATCGGTTGCTCTTACAAACATGTTACCGAGGCAAGGTCTTC 480
Db 2343 GGGCCTTTCTTCTGTACGACAGGATCGCCAGCACCACTGTTACAGGGGGGAGGGGTTC 2402
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAAATAAGACAGTTTCAAGAAATGATTTTCT 540
Db 2403 ACCGAGGCAACATCGCGCCATGATCGTTAA CAGACCGTGCACAGATGATCTTCAGC 2462
QY 541 AGCAAGGACAGGTTATCGTCATGAATCTTGACCTTCCACCAATAAATATTTGACAGAC 600
Db 2463 AGCGAGGCGCAGGGCTACAGGCACATGAA CCTGACACAGCACCAACAAGTACTGGACCA 2522
QY 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATATCACTCC 660
Db 2523 AACAAACGACCCACAGACCAACGACCGGCTGCTTCGGCGCCCTGCGAGGAGTACACAGC 2582
QY 661 ACRAACAATCAACATCCCTTCCATCTCTTAAACCTTCACTCCCTCCGCGCCACAGTAACCTCG 720
Db 2583 ACCAAGAACAGACCTTGGCCCCCAGCAAGATGCCAGGCCCTTCCACCGCCAGGGCCC 2642
QY 721 AGCATTCACCTTACAAATCTCAAAATTAATACTGTCTAAATCTGGAACCTATGAACCCCAAGT 780
Db 2643 GAGATCAAGCCCAACAGCACCCCAACGACCGCCACCTTGAACACACCGGACCCCAAC 2702
QY 781 AGGACGATGAGGACCTTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGGCCCCACACA 840
Db 2703 AACGACGACGAGGACCTTGATCAC CAGCGCGAGCGCGCGAGGAGGAGGAGCCCTACACC 2762
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTTGTCACCTCTCTTCACTA 900
Db 2763 ACCAGGACCGCGTGACCAAGAGGGGCTGAGCAGGACCACTGCTCTTACCTTAGCCCT 2822
QY 901 CATCCAAGCACCTCACAAACATGAGCAAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
```

```
Db 2823 CAGCCACGACCCCTCAGCAGGAGGCAACACACCGACACAGCGGACCGTACC 2882
Qy 961 GAGCAAAATGGAACCGACCCCAACAAACAACAGCAACGCTCTCTCAACAATACTAATA 1020
Db 2883 GAGCCCAACAAGACCAACACACCGCCAGCCAGCATGCTCTCTCAACAACACCGCC 2942
Qy 1021 ACTCCACCTATACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCACCGCAAC 1080
Db 2943 ATCAGCAACAACAACAGCAAGAAACAATCTCAGCACCTTGAGCGTGGAGT 3002
Qy 1081 ATCACAATAATGATACACAACGTAAGTACAGCAAGAAACGCAACCAATGCTCAGTTG 1140
Db 3003 ACCACCACTAGCAACCAAGCAACCGCCACCGAGACGAGCAGCAGCGCCCTAGC 3062
Qy 1141 AACACAATCTAGATCCAAAGAAATCCCAACAGCAAGCAACCAACAGCAACCAAC 1200
Db 3063 AAGACCACTCTCCCTCCACCGCAACCTGACCAACCGCAAGAGCAGCAACCAACCAAG 3122
Qy 1201 AACATCATGAGCAGCATAGATATACAGCAACCAACCGCCACCAATCTCTCCGAT 1260
Db 3123 GGGCCACCAACCGCCCTTAACATGACCAACCGCCACCTGACAGCGCCACCGCCAC 3182
Qy 1261 TCTAGTCCGCAACCGCCCTCTATATACATTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 3183 CCCAACCCCAACCAACGACCTGCTGATCTTCAGGAAGAGAGGAGCATCTCTGGAGG 3242
Qy 1321 GAAGCGCATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAATGATTTTGACCCA 1380
Db 3243 GAGGCGATATGTTCCCTTCTCGACGCGCTGATCAACGCGCTTATCGACTGACGCC 3302
Qy 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCTCTAGTCTGCTGCTCGCTGAGGAA 1440
Db 3303 GTGCCCAACACCAAGCACTCTTCAGCAGAGCAGCAGCGCGCGCGCGCGAGAG 3362
Qy 1441 GATCAACATGCTCTCCCAATATAGTTTAACTTTATCTTATTTCTTAATAATAAGAG 1500
Db 3363 GACCAGCAGCGCGCCCAACATCAGCTGACCTGACCTTCTCCCAACATCAACGAG 3422
Qy 1501 AACCTGCTCTCTGAGAGAAATCAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 1560
Db 3423 AACACCGCTCAGCGCGAGACGAGAACGACTGCGAGCGGAGCTGAGGATCTGGAGC 3482
Qy 1561 GTTACAGGAGTACCTCGCGCGAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAAT 1620
Db 3483 GTGCAGGAGGACGCTCGCGCGCGCTGAGCTGGATTCCTCTTCGCGCCCGCGCATC 3542
Qy 1621 GAAGCACTTACACTGCTGTTTTTAATAAATCAAAATTTGGTCTCGAGGTTGAGG 1680
Db 3543 GAGGCGCTGTACACCGCGCGCTGATCAAGAACCAAGAACCACTGGTGTGAGGCTGAGG 3602
Qy 1681 CGCTAGCGCAATCAAACTGCGCAATCTTTGGAATCTTTATGAGAGTCACAACTGAGGA 1740
Db 3603 AGGCTGGCCACCAAGCGCGCAAGAGCTGAGCTGCTGAGGCTGACCAACCGAGAG 3662
Qy 1741 AGAACATCTCTCTTAATCAATAGACATGCTATGATCTTTCTACACAAAGATGGGAGGA 1800
Db 3663 AGGACCTTCAGCTGATCAACAGGCAACGCTGCTGCTGACCGAGTGGCGCGC 3722
Qy 1801 ACATCGAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTGTCCAAATAAT 1860
Db 3723 ACCTGCAAGGTCTGGGCGCGACTGCTGATCGGCATCGAGGACCTGAGCAGGAACATC 3782
Qy 1861 TCAGAGCAAAATGACCAATTAAGAGGCAACAAAGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Db 3783 AGCGAGCAGATCGACAGATCAAGAGGAGCAGAGCAGAGAGGAGGAGGAGGAGGAGGAG 3842
Qy 1921 GGTGTAATGTTGGACATCCGACTG 1946
Db 3843 GCGCGCAAGTGTGGACGCGACTG 3868
```

RESULT 12

```
US-10-066-506A-3
; Sequence 3, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION: Case C.
; APPLICANT: Grogan, Michael C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506A-3
```

Query Match 38.1%; Score 780; DB 15; Length 1841;

Query Local Similarity 74.8%; Pred. No. 3.4e-216; Mismatches 0; Indels 0; Gaps 0;

Matches 978; Conservative 0;

```
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTTTAATCCAAAGTATAAAAATCTCCCT 60
Db 1 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGGAGCAAAAAATCTCCCC 60
Qy 61 GTTTTGAATTTGCTAGTAACAGCAACCTCAAGATTAGATTAGTGTGCTCCGGAAC 120
Db 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCTATGCTCCGGAAT 120
Qy 121 CTCCAAAGACAGAGATGTTCTATCTGATGGGATTTACACTGAGTGGSCAAAAGTTTCT 180
Db 121 CTCAGAAGACAGAGAGCTGCTATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTTCT 180
Qy 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTGAGGCATCCAAGGATGGCTTTTCAGGACAGGTGTTACCTCCCAAGAT 240
Qy 241 GTTCAGTATACGGAAGGAGAGAAAGCCAAACATGTTTCAATATAGTGTACAGACCT 300
Db 241 GTTCAGTACACAGAGGGGAGAGCAAAACATGCTACATATAAGTGTACAGGATCCC 300
Qy 301 TCTGGAATTCCTTGTCTGCTGGATCCTCCAGTATATATCCGATTAACCTTAAATGTAA 360
Db 301 TCTGGAATTCCTTGTCTGTTAGATCCTCTTACCAACATCCGTAATCCGAAATGCAAA 360
Qy 361 ACTGTTATCATATTTCAAGGTCAAAACCTCATGACAGGGGATTCCTCCATTTGTGG 420
Db 361 ACTATCATATATTTCAAGGTCAAAACCTCATGACAGGGGATTCCTCCATTTATGG 420
Qy 421 GGGCATTTTTTCTGTATGATCGGTTGCTCTTCAACAATGTACCGAGGCAAGGTCTTC 480
Db 421 GGAGATTTTTTCTGTATGATCGCATTCCTCTCCCAACAATGTACCGAGGCAAGGTCTTC 480
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTACAGATGATTTTTTCT 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTACAGATGATTTTTTCT 540
Qy 541 AGGCAAGGACAGGTATTCGTACATGAACCTTACCTCCCAACAATATAATTTGGACAAGC 600
Db 541 CGGCAAGGACAGGTATTCGTACATGAATGATCTGATCTTACTATATAATTTGGACAAGT 600
Qy 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTTGGCATCTCCCAAGATACAACTCC 660
Db 601 AGTAACGGAACGCAACGAATGACACTGGAATGTTTTCGGCGCTCTTCAAGAAATACAATCT 660
Qy 661 ACAACAAATCAAAATGATGCTTCTTAAACCTTCAATCCCTGCCCAACAGTAATCTCG 720
```

[illegible]

RESULT 13
US-10-066-506A-1
; Sequence 1, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 1 and Marburg virus strain Musoke Glycoprotein 2
; OTHER INFORMATION: and Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-1

```
Query Match      36.5%; Score 746.6; DB 15; Length 2252;
Best Local Similarity 99.5%; Pred. No. 2.1e-206;
Matches 749; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy	1294	AGAAAGAAACGATCGATCCTCTCGAGGGAAGGCGACATGTTCCGTTTCTCGAATGGGTTA	1353
Db	1492		
		AGAACTCGACGATCGATCCTCTCGAGGGAAGGCGACATGTTCCGTTTCTCGAATGGGTTA	1551
Qy	1354	ATAAATGCTCCAAATTGATTTTGGACCCAGTTTCCAAATACAAAAACAATCTTTGATGAATCC	1413
Db	1552		
		ATAAATGCTCCAAATTGATTTTGGACCCAGTTTCCAAATACAAAAACAATCTTTGATGAATCC	1611
Qy	1414	TCTAGTTCCTGGTGCTCGGCTGAGGAAGATCAACATGCGCTCCGCCAATAATTAGTTTAACT	1473
Db	1612		
		TCTAGTTCCTGGTGCTCGGCTGAGGAAGATCAACATGCGCTCCGCCAATAATTAGTTTAACT	1671
Qy	1474	TTATCTATTTTCTTAATATAAATGAGAAACAATGCGCTACTCTGAGAGAAAATGAGAATGAT	1533
Db	1672		
		TTATCTATTTTCTTAATATAAATGAGAAACAATGCGCTACTCTGAGAGAAAATGAGAATGAT	1731
Qy	1534	TGTGATGACAGATTAGAAATTTGGAGCGTTTCAGAGGATGACCTGGCCGCGACAGGCTCAGT	1593
Db	1732		
		TGTGATGACAGATTAGAAATTTGGAGCGTTTCAGAGGATGACCTGGCCGCGAGGCTCAGT	1791
Qy	1594	TGGATACCGTTTTTTTGGCCCTGGAAATGGAAGGACTTTTACACTGCTGTTTTTAAATTTAAAAAT	1653
Db	1792		
		TGGATACCGTTTTTTTGGCCCTGGAAATGGAAGGACTTTTACACTGCTGTTTTTAAATTTAAAAAT	1851
Qy	1654	CAAAAACAATTTGCTGTGACGGTTGAGCGGCTATAGCCAAATCAAACTGCGCAAAATCCTTGGAA	1713
Db	1852		
		CAAAAACAATTTGCTGTGACGGTTGAGCGGCTATAGCCAAATCAAACTGCGCAAAATCCTTGGAA	1911
Qy	1714	CTCTTATTTGAGAGTCAACACTGAGGGAAGAACATTTCTCTTAATCAATAGACATGCTATT	1773
Db	1912		
		CTCTTATTTGAGAGTCAACACTGAGGGAAGAACATTTCTCTTAATCAATAGACATGCTATT	1971
Qy	1774	GACTTTCTACTCAAGATGGGAGGAAACAATGCAAGTGTCTGGACCTGATTTGTCATC	1833
Db	1972		
		GACTTTCTACTCAAGATGGGAGGAAACAATGCAAGTGTCTGGACCTGATTTGTCATC	2031
Qy	1834	GGGATAGAGAATTGTGTCAAAAATAATTTAGAGCAAAATTCGCAAAATTAATAAGGACGAA	1893
Db	2032		
		GGGATAGAGAATTGTGTCAAAAATAATTTAGAGCAAAATTCGCAAAATTAATAAGGACGAA	2091
Qy	1894	CAAAAAGGGGACTGGTTGGGGCTCTGGGTGGTAAATGGTGGACATCCGACTCGGGGT	1953
Db	2092		
		CAAAAAGGGGACTGGTTGGGGCTCTGGGTGGTAAATGGTGGACATCCGACTCGGGGT	2151
Qy	1954	CTTACTAACTTGGGCATTTTGCTACTATTATCAATAGCTGCTGTGATGCTCTATCCTGT	2013
Db	2152		
		CTTACTAACTTGGGCATTTTGCTACTATTATCAATAGCTGCTGTGATGCTCTATCCTGT	2211
Qy	2014	ATTGTGCTATCTTTACTAAATATATCGGATAA	2046
Db	2212		
		ATTGTGCTATCTTTACTAAATATATCGGATAA	2244

```

RESULT 14
US-10-353-856-9
; Sequence 9, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawakoa, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18890
; TYPE: DNA

```

